Strategies for engineering sensory photoreceptor chimeras

Strategien zur Entwicklung sensorischer Photorezeptorchimären

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biophysik

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Robert Ohlendorf, M. Sc.

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Lebenswissenschaftlichen Fakultät
Prof. Dr. Richard Lucius

Gutachter:
1. Prof. Dr. Andreas Möglich
2. Prof. Dr. Keith Moffat
3. Prof. Dr. Andreas Herrmann

Tag der mündlichen Prüfung: 11.03.2016

∇

1 Abstract

Sensory photoreceptors mediate diverse responses to ambient light in all domains of life. As many other proteins involved in signal transduction, sensory photoreceptors often have a modular architecture. Different modules enable light perception (sensor module) and biological output function (effector module) and are often structurally and functionally coupled by α-helical linkers. Optogenetics employs these light-regulated proteins for controlling cellular processes with high spatiotemporal accuracy and minimal invasiveness. Rewiring sensor and effector modules from different proteins yields photoreceptor chimeras to put target cellular signaling pathways under light control and e.g. dissect their underlying network architecture. Thereby, the major challenge is fusing the linkers from both modules in a way that preserves signal transduction within the chimera.

This study explores strategies for efficiently engineering photoreceptor chimeras. The first part followed a rational-design approach guided by sequence and structure homology of the parent proteins. It focused on cyanobacteriochrome (CBCR) photosensors that respond to diverse light colors and can be switched between two signaling states using light of different wavelengths. Together with their small size, these properties make CBCRs attractive building bricks for optogenetic tools. Starting out from the well-characterized, light-repressed histidine kinase YF1, its blue-light sensor was replaced by CBCR sensor modules. While most chimeras did not show any light-regulated kinase activity in *Escherichia coli*, one functional chimera, CF, exhibited 4.6-fold induction by green and repression by red light. In a next step, the same sensor was rewired to a different effector module – an adenylate cyclase that is originally regulated by homologous sensor modules. However, the respective rationally designed CCyc chimera did not respond to light in bacterial reporter assays.

Overall, the small number of functional chimeras and the low predictability of the rational-design outcome led to the development of a novel brute-force approach, termed PATCHY (primer-aided truncation for the creation of hybrid enzymes). Complementary to rational design, this method generates libraries of fusion variants with all desired combinations of parent linker sequences. Bacterial reporter assays then allow screening these libraries with high throughput and isolating light-regulated chimeras. In comparison to previous, comparable approaches PATCHY is quick and easy to use, utilizes materials readily available in most labs and produces uniform distributions of fusion variants.

Fusion libraries of the *Bacillus subtilis* YtvA-LOV (light-oxygen-voltage) sensor and the *Bradyrhizobium japonicum* FixL histidine kinase yielded numerous blue-light regulated chimeras featuring different linker lengths and sequences. Linker variation produced light-induced and light-repressed chimeras, each group complying with a heptad periodicity of linker lengths (7n+3 and 7n+2 residues). With less than 5% of all possible variants exhibiting light regulation, the fine-tuning of linker sequence and protein function becomes evident. Establishing a bacterial reporter assay for adenylate cyclases allowed screening similar fusion libraries of *Bs*YtvA-LOV and the adenylate cyclase bPAC. Resulting light-regulated variants showed low dynamic ranges and might indicate little compatibility of these modules.

Hence, the PATCHY strategy for scanning linker variants between two modules does not only facilitate the development of functional photoreceptor chimeras for cellular manipulation, e.g. in

optogenetics. It also provides information on the relation between linker properties and protein function, and thereby contributes to a deeper mechanistic understanding of signal transduction within modular signal receptors.

2 Zusammenfassung

Sensorische Photorezeptoren vermitteln vielfältige Lichtreaktionen in allen Domänen des Lebens. Wie zahlreiche andere Signaltransduktionsproteine, sind sie modular aufgebaut. Dabei dienen verschiedene, durch α-helikale Verbindungsstücke (‚Linker') strukturell und funktional gekoppelte, Module der Lichtperzeption (Sensormodul) und der Erzeugung einer biologischen Aktivität (Effektormodul). In der Optogenetik dienen diese lichtregulierten Proteine der sowohl zeitlich und räumlich präzisen, als auch minimalinvasiven Kontrolle diverser zellulärer Prozesse. Die Rekombination von Sensor- und Effektormodulen unterschiedlicher Elternproteine erzeugt Chimären, die eine gezielte Lichtregulation ausgesuchter Signalwege ermöglicht, z.B. zur Aufklärung deren Netzwerkstruktur. Die größte Herausforderung besteht dabei darin die Linker beider Module in einer Weise zu fusionieren, die eine Signalweiterleitung zwischen Sensor und Effektor ermöglicht.

Die vorliegende Arbeit erforscht Strategien zum effizienten Bau chimärer Photorezeptorproteine. Ein erster Abschnitt folgte einem rationalen, auf Sequenz- und Strukturhomologie der parentalen Proteine basierenden, Ansatz. Im Fokus standen hierbei Photosensoren aus der Familie der Cyanobakteriochrome (CBCR). Diese reagieren auf diverse Spektralbereiche des sichtbaren Lichtes und können mit unterschiedlichen Wellenlängen zwischen zwei Signalzuständen geschaltet werden. Zusammen mit der geringen molekularen Größe, machen diese Eigenschaften CBCR zu attraktiven Bausteinen optogenetischer Werkzeuge. Ausgehend von der gut charakterisierten, lichtreprimierten Histidinkinase YF1, wurde deren Blaulichtsensor durch CBCR-Sensormodule ersetzt. Während der Großteil aller erzeugten Chimären im Testsystem *Escherichia coli* keine lichtabhängige Kinaseaktivität aufwies, zeigte das Konstrukt CF 4,6-fache Induktion durch grünes Licht bzw. Repression durch Rotlicht. Im nächsten Schritt sollte das gleiche Sensormodul zur Lichtkontrolle einer Adenylatzyklase verwendet werden welche ursprünglich durch strukturell homologe Sensoren reguliert wird. Das resultierende, rational geplante Fusionsprotein CCyc zeigte jedoch nur lichtunabhängige Funktion.

Die insgesamt geringe Ausbeute funktionaler Chimären im rationalen Ansatzes und die schlechte Vorhersagbarkeit dessen Erfolges führten zur Entwicklung einer neuen ‚Brute-Force'-Methode namens PATCHY (primer-aided truncation for the creation of hybrid enzymes). Komplementär zum rationalen Ansatz generiert PATCHY Bibliotheken von Fusionsvarianten aller Linkerkombinationen. Bakterielle Testsysteme erlauben danach die schnelle Dursuchung der Bibliotheken nach funktionalen, lichtregulierten Chimären. Im Vergleich zu bisher publizierten Methoden ist PATCHY schnell und leicht anzuwenden, benötigt im Labor gewöhnlich vorrätige Materialien und produziert uniforme Verteilungen von Fusionsvarianten.

Entsprechende Bibliotheken mit Fusionen des Sensors *Bacillus subtilis* YtvA-LOV (light-oxygen-voltage) und der Histidinkinase *Bradyrhizobium japonicum* FixL enthielten zahlreiche blaulichtregulierte Chimären mit unterschiedlichen Linkerlängen und -sequenzen. Dabei führte die Linkervariation sowohl zu lichtaktivierten, als auch zu lichtreprrimierten Konstrukten. Die Linkerlängen beider Klassen folgen einer Heptadenperiodizität und bestehen bevorzugt aus 7n+3 bzw. 7n+2 Aminosäuren. Weiterhin verdeutlicht die Tatsache, dass weniger als 5% aller möglichen Fusionen als lichtreguliert identifiziert werden konnten die feine Abstimmung von Linkersequenz und Proteinfunktion. Etablierung eines bakteriellen Testsystems für

Adenylatzyklasen ermöglichte die Analyse ähnlicher Fusionsbibliotheken aus *Bs*YtvA-LOV und der Adenylatzyklase bPAC. Der geringe Aktivierungsfaktor der identifizierten, lichtregulierten Varianten könnte hier auf eine eingeschränkte Kompatibilität der beiden Module hinweisen.
Die Anwendung von PATCHY für die systematische Analyse von Linkervarianten zweier Module trägt daher zur Entwicklung chimärer Rezeptorproteine für die Manipulation zellulärer Prozesse, z.B. in der Optogenetik, bei. Darüber hinaus gibt die funktionelle Information aus einer Vielzahl von Fusionsvarianten aber auch Aufschluss über die Verbindung zwischen den Eigenschaften des Linkers und der Proteinfunktion bzw. der Signaltransduktion innerhalb des Rezeptorproteins.

3 Introduction

3.1 Sensory photoreceptors

Throughout nature light either serves as energy source, like in photosynthesis, or transports information influencing behavior or physiology in organisms of all kingdoms. Examples for adaptation to environmental light conditions include phototropism in higher plants[1], phototaxis in certain flagellates[2] and visual perception in the animal eye[3]. Sensory photoreceptors mediate the transformation of the electromagnetic light signal into a biological function and often show a modular architecture[4] in common with most other signal-transduction proteins. Separate modules, usually implemented as protein domains, are associated with signal perception (sensor module) and exertion of a biological function (effector module). Sensor modules absorb visual light via the conjugated π-electron system of an organic chromophore that derives either from amino-acid side chains or small metabolites like flavins, bilins or retinals. The photon energy drives the photocycle – reversible photochemical reactions of the chromophore, which in turn induce conformational changes within the sensor module. The light-induced structural changes close to the chromophore then propagate to the effector module and modulate its activity.

Based on their chromophores and the respective photochemistry, photoreceptors fall into classes with various spectral sensitivities[5] (Figure 1). UV-B (ultraviolet B) receptors, like UVR8 (UV-B resistance 8) from *Arabidopsis thaliana*, use a triad of tryptophan side chains to absorb UV-B light. Cryptochromes, LOV (light-oxygen-voltage) and BLUF (blue-light sensors using flavin adenine dinucleotides) proteins bind flavin chromophores and are sensitive to blue light. While all these classes perceive rather narrow windows of the electromagnetic spectrum, rhodopsins tune their retinal chromophores to various sensitivities from UV to the red range. Plant phytochromes use linear tetrapyrrols (bilins) to respond to red and near-infrared light, whereas their algal homologues span the entire visual spectrum[6]. A similar spectral diversity occurs in cyanobacteriochromes (CBCR) that also use bilin chromophores. Apart from rhodopsins all other

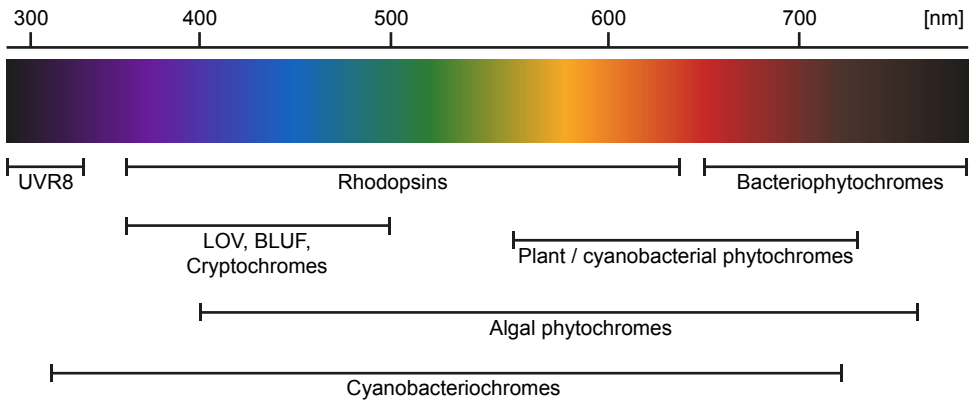

Figure 1: Spectral sensitivities of photoreceptor proteins.

sensory photoreceptors are cytosolic proteins. The following sections introduce LOV and CBCR photoreceptors, which are central to the present study.

3.1.1 Light-Oxygen-Voltage photoreceptors

LOV photoreceptors are widely distributed among bacteria, archaea, plants and fungi where they mediate diverse processes like phototropism, circadian and developmental phenomena, stress responses and virulence[7]. At the structural level LOV domains constitute a subclass of the Per-ARNT-Sim (PAS) family, which include versatile sensor and interaction modules of signaling proteins[8,9]. LOV domains comprise a five-stranded antiparallel β sheet and four α helices[9,10] with the flavin cofactor bound in a cleft formed by the central β sheet and two adjacent helices[11] (Figure 2a).

A characteristic photochemistry (Figure 2b) distinguishes LOV sensors from other flavin-based photoreceptors[11,12]. After photon absorption at wavelengths around 450 nm by the dark-adapted state (D_{450}), the flavin mononucleotide (FMN) chromophore either fluoresces or enters the triplet L_{660} state by highly efficient intersystem crossing within nanoseconds. Subsequently, the system returns to the ground state via internal conversion or enters the S_{390} state by formation of a thioether bond between atom C(4a) of the flavin cofactor and a highly conserved, proximal cysteine residue[13,14]. Reduction of the flavin in the photoadduct increases the pKa of N5 and elicits its protonation[15]. As a consequence, the side chain of a nearby glutamine residue rotates by

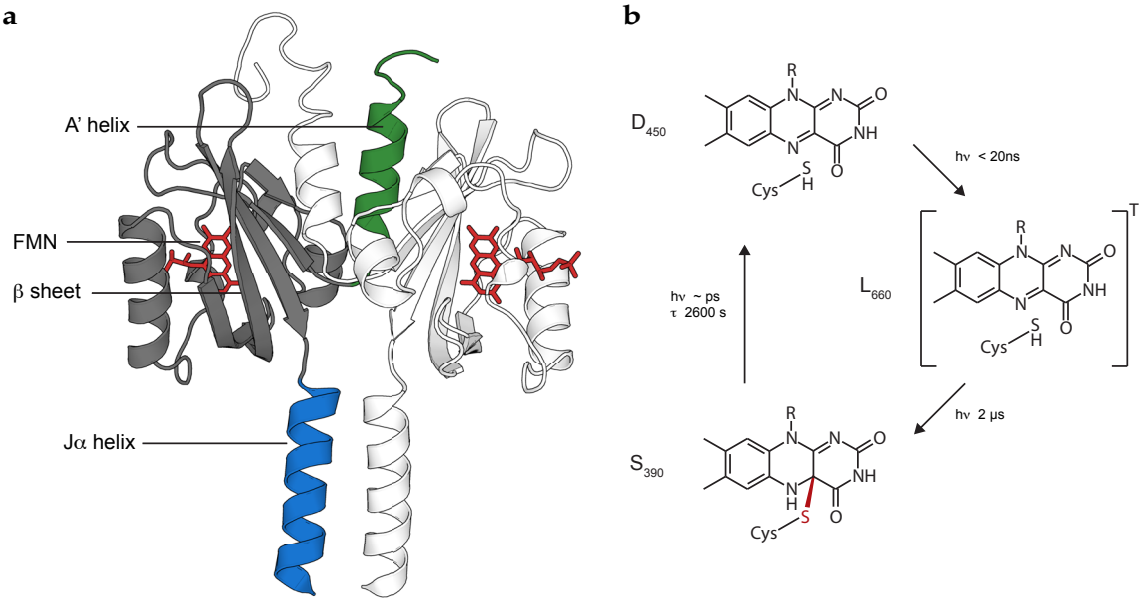

Figure 2: LOV domains. (**a**) The crystal structure of the *Bs*YtvA-LOV domain (residues 1-143, PDB: 4GCZ) shows a head-to-head homodimer, each monomer incorporating a flavin chromophore (red). N- and C-terminal α helices (green and blue) form coiled coils at the dimer interface. (**b**) The LOV photocycle (adapted from Möglich, Yang *et al.*, 2010). After photon absorption by the flavin at around 450 nm the dark state (D_{450}) enters the L_{660} triplet state via intersystem crossing. L_{660} converts to the signal state S_{390} by formation of a thioether bond between the flavin and a proximal cysteine residue. The system thermally relaxes to the ground state.

180° inducing rearrangements in hydrogen bonding that propagate through the LOV domain[16–18], which alter structure and dynamics of the N- and C-terminal helices (Figure 2a). Recent data suggest that in the absence of this cysteine, reduction of the flavin by formation of a neutral semiquinone state and N5 protonation suffice to trigger downstream signaling[19]. The photoadduct thermally reverts to the resting state, with time constants greatly varying among members of the LOV family. While phototropin LOV domains recover within 10-100 seconds[20] this process takes several hours in LovK of *Caulobacter crescentus*[21]. In general, the recovery kinetics strongly depend on the protein context and differ between isolated sensors and full-length proteins[22]. For the LOV2 domain of phy3 from *Adiantum*, illumination with near-UV light is reported to facilitate an ultrafast regeneration within picoseconds[23].
Albeit the photochemical reactions are closely similar in all LOV domains the resulting signal propagation within the protein is not conserved[7]. However, the N- and C-terminal helices, linking LOV domains to the rest of the protein, play a major role. Structural studies on *Avena sativa* phototropin 1 LOV2 (*As*LOV2) demonstrated partial light-induced unfolding of the Jα helix[24], whereas in the VIVID protein from *Neurospora crassa* an N-terminal, helical extension is reoriented upon illumination allowing dimerization[17].

The *Bs*YtvA-LOV domain that is central to this study represents another signaling paradigm. YtvA is involved in the blue-light stress response of *Bacillus subtilis*[25] and comprises an N-terminal LOV domain adjacent to a C-terminal STAS domain (sulfate transporter/anti-sigma-factor-antagonist). Independent of illumination the YtvA protein[26], as well as the isolated LOV sensor[10], form constitutive homodimers[26] as found in the crystal structure of the YtvA-LOV domain[10] (Figure 2a). Upon light activation $YtvA_{447}$ undergoes the characteristic photochemistry described above, reaching the covalent adduct ($YtvA_{390}$) via the FMN triplet state ($YtvaA_{660}$) within 2 μs[27]. In the absence of blue light the $YtvA_{447}$ dark state recovers within 2600 s at room temperature, while recovery of the isolated YtvA-LOV domain takes 3900 s[28]. However, UV light application drives photoconversion to the dark state with moderate yields within picoseconds[29]. In contrast to other LOV domains like *As*LOV2, YtvA-LOV features a more polar, C-terminal Jα helix[30,31], which does not pack against the LOV-domain core but points outwards, thereby exposing hydrophobic residues on the outer surface of the β-sheet (Figure 2a). This presumably accounts for dimerization of the YtvA-LOV domain[10]. For signal propagation within the YtvA dimer Möglich *et al.* proposed a reorientation of the YtvA-LOV monomers, which is transmitted through the Jα helices to the effector modules[32,33].

3.1.2 Cyanobacteriochromes and other bilin photoreceptors

Bilin photoreceptors comprise phytochromes and cyanobacteriochromes and covalently bind a linear tetrapyrrol (bilin) chromophore via a conserved cysteine residue. While phytochromes are wide-spread among different bacteria, diatoms, fungi, algae and plants (Table 1), cyanobacteriochromes exclusively occur in cyanobacteria where they govern photochromic and phototatic responses[34–36]. Both photoreceptor classes share a mutual photochemistry that involves isomerization of the chromophore between *15Z* and *15E* configurations associated with rotation

of the terminal bilin D-ring[37–39] (Figure 3a). These two semi-stable states possess distinct absorption properties (photochromism) and interconvert within milliseconds by application of different light qualities. After illumination the thermodynamically more stable state (in the following always mentioned first) also recovers thermally within seconds to hours. Whereas

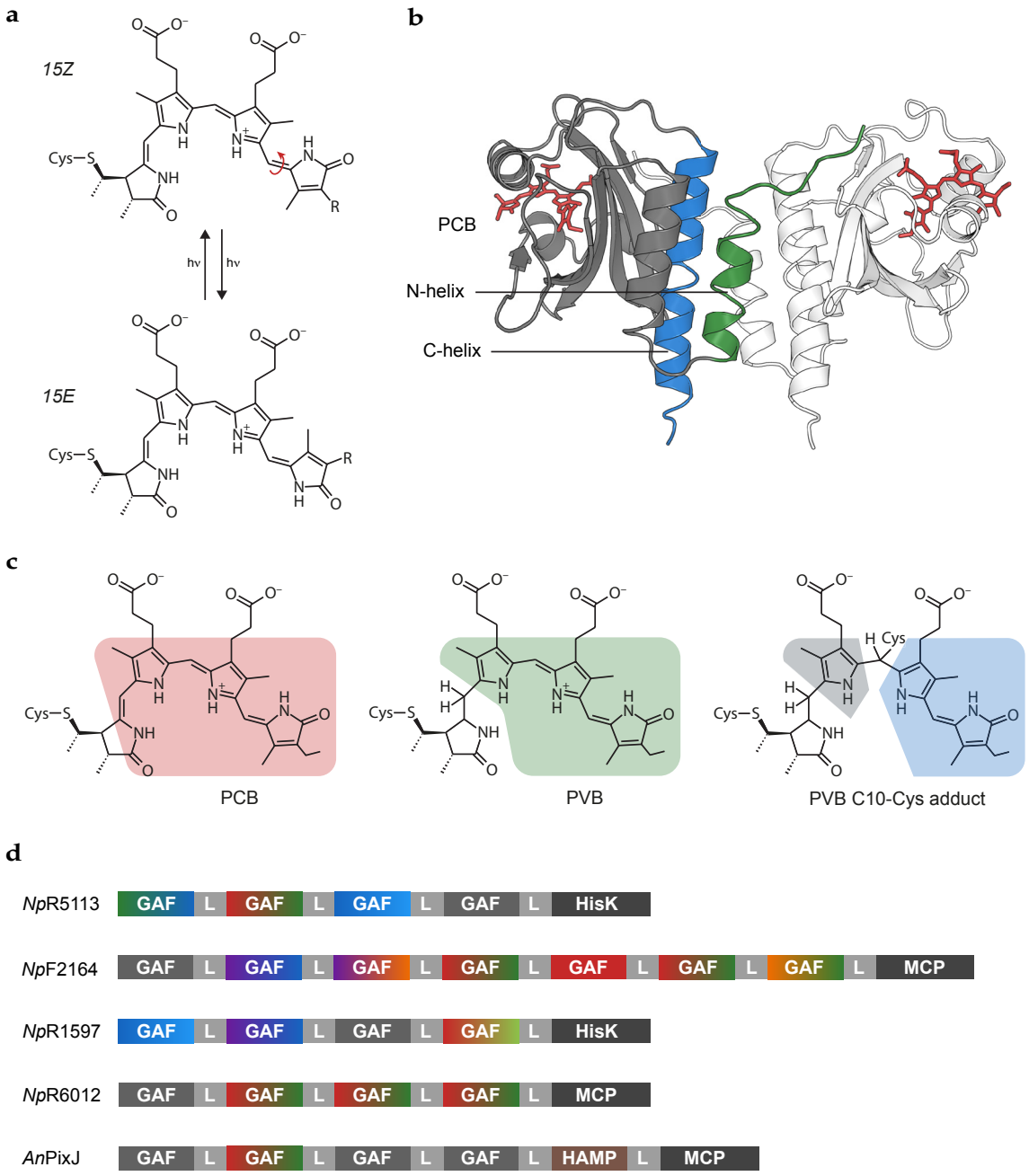

Figure 3: CBCR photoreceptors. (**a**) Photochemistry of CBCRs. The bilin chromophore is covalently bound to the protein via a conserved cysteine residue. Different wavelengths induce isomerization between *15Z* and *15E* isoforms inducing a rotation of the distal D ring. (**b**) Structure of CBCR sensors (*An*PixJ residues 12-189, PDB: 3W2Z). A typical GAF fold harbors the PCB chromophore (red) and features N- and C-terminal helices positioned at the interface of the head-to-head homodimer. (**c**) Delocalized π-electron systems of PCB and PVB chromophores. Reversible linkage to a second cysteine in the protein splits the system and shifts the chromophore absorption to shorter wavelengths (adapted from Rockwell *et al.*, 2011) (**d**) Domain architectures of selected CBCRs show that multiple sensors usually connect to a single, C-terminal effector. 'L' denotes linker regions between the domains.

canonical phytochromes switch between red-absorbing (Pr) and far-red-absorbing states (Pfr), cyanobacteriochromes show a vast diversity of photocycles including blue/UV-, violet/orange-, or red/green-sensitive types. Recently found phytochromes from algae show a similar spectral versatility[6].

While all cyanobacteriochromes initially incorporate the same red-absorbing phycocyanobilin (PCB) chromophore (Figure 3a), several mechanisms evolved for tuning its sensitivity to detect the entire visible spectrum[40] (Table 1). Green-red (G/R) CBCRs have a green-absorbing *15Z* dark or ground state and a red-absorbing *15E* photoproduct[35]. In a unique 'protochromic' photocycle three conserved residues facilitate reversible protonation of the green-absorbing *15Z* state, which shifts the absorption to a red-absorbing *15E* state[41]. Red-green (R/G) CBCRs feature an opposite photocycle with a red-absorbing *15Z* ground state and a *15E* state, where conserved phenylalanine residues twist the chromophore geometry and shift its absorption into the green range[42–44]. Several subclasses trade under the term Dual-Cys CBCRs and feature a second cysteine in addition to the canonical cysteine that permanently fixes the bilin to the protein moiety. DXCF CBCRs, where the additional cysteine is situated in the eponymous aspartate-x-cysteine-phenylalanine motif, autocatalytically isomerize PCB to phycoviolobilin (PVB)[45–47] (Figure 3c) and usually have blue-absorbing *15Z* ground states and green-absorbing *15E* photoproducts[40,48]. The second cysteine reversibly forms a covalent bond to the chromophore in the ground state splitting

Table 1: Bilin photoreceptors (adapted from Ziegler and Möglich, 2015)[4]. Species abbreviations: *Am, Acaryochloris marina; An, Anabaena sp.; Cp, Cyanophora paradoxa; Dt, Dolichomastix tenuilepis; Es, Ectocarpus siliculosus; Np, Nostoc punctiforme; Te, Thermosynechococcus elongatus*.

Class	Photoreceptor	Sensor	Chromophore	Absorption *15Z*/*15E* [nm]	Paradigm
Phytochromes	Plant	PAS-GAF-PHY	PCB, PΦB	650-670/700-730	*At*PhyB
	Bacterial/ fungal	PAS-GAF-PHY	BV	700/750	*Dr*BPhy
		PAS-GAF-PHY	BV	750/700	*Pa*BPhy
	Cyanobacterial	PAS-GAF-PHY	PCB	660/705	*Sy*Cph1
	Algal	PAS-GAF-PHY	PCB	595/725	*Dt*Phy1
		PAS-GAF-PHY	PCB	440/635	*Cp*GPS1
		PAS-GAF-PHY	PΦB	690/565	*Es*PHL1
	Cph2	GAF-PHY	PCB, PΦB	643/690, 655/700	*Sy*Cph2[133]
Cyanobacteriochromes	G/R	GAF	PCB	535/670	*Sy*CcaS
	R/G	GAF	PCB	650/545	*An*PixJ
	DXCF	GAF	PCB/PVB	430/530	*Te*PixJ
		GAF	PCB	435/600	*Np*F4973
	Insert-Cys	GAF	PCB	400/590	*Np*F2164
	New Dual-Cys	GAF	PCB	640/415	*Am*AM1_1186

the conjugated π-electron system (Figure 3c) causing the blue-absorbing *15Z* configuration[42,48,49]. Insert-Cys CBCRs harbor the second cysteine, but no DXCF motif and use a similar mechanism to tune the absorption of their PCB chromophores to UV-blue and blue-orange photocycles[40,50]. Finally, AM1_1186 from *Acaryochloris marina* represents a novel third type of Dual-Cys CBCRs, where the second cysteine is situated at a unique position. Closely related to R/G-CBCRs the PCB chromophore features a red-absorbing *15Z* ground state, but reversible linkage of the second cysteine drastically shifts the absorption of the *15E* state to the blue[51]. Comparable protein-chromophore interactions might tune the chromophore absorption in algal phytochromes.

The photosensor module of phytochromes comprises the three domains PAS, GAF (cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA) and PHY (phytochrome specific), which together coordinate the chromophore. Cph2 sensors retain GAF and PHY[36], whereas in cyanobacteriochromes a single GAF domain suffices for chromophore binding and reversible photochemistry[40] (Table 1). Still, multiple CBCR sensors often occur in tandem in a single protein allowing signal integration at the C-terminal effector module[52] (Figure 3d). Structural data of the R/G-CBCR *An*PixJ-GAF2[42] from *Anabaena sp.* (Figure 3b) and DXCF-CBCR *Te*PixJ-GAF[42,49] from *Thermosynechococcus elongatus* show the typical GAF fold with the domain core comprising an antiparallel beta-sheet and several short helices. In the crystal cell both CBCRs form parallel homodimers with N- and C-terminal linker helices forming a central bundle[42]. The same data suggest changes in hydrogen bonding between the chromophore and an aspartate residue upon photoisomerization[42]. Hence, like in LOV domains, local rearrangement in hydrogen bonding at the chromophore may propagate across the sensor module to trigger downstream signaling[42] via the N- and C-terminal helices.

3.2 Allostery of sensory photoreceptors

Light absorption triggers a series of events within a photoreceptor protein. The previous section described examples of how the initial photochemistry alters the interaction between the chromophore and proximal protein residues and modifies hydrogen bond networks within the sensor module (3.1). While these processes are largely investigated for most photoreceptors, subsequent conformational changes of the sensor module and the signal propagation to distal effector modules are still obscure. Receptor proteins exist in equilibrium between at least two states with different biological activities. The light signal shifts the equilibrium between these states and alters the net activity of the protein[33]. Thereby, the mechanisms, which allosterically couple light absorption in the sensor to the function of the effector module, are diverse.

A recent publication by Ziegler and Möglich distinguishes associating and non-associating types (Figure 4) depending on whether photoreceptors change their oligomeric state upon light absorption[5]. For example VIVID from *Neurospora crassa* undergoes light-induced homodimerization[17] (Figure 4a), whereas homodimers of UVR8 from *Arabidopsis thaliana* dissociate upon UV illumination[53,54]. Plant phytochromes PhyA and PhyB from *Arabidopsis thaliana* form heterodimers (Figure 4b) with light-inert PIF3 or PIF6 proteins under red light,

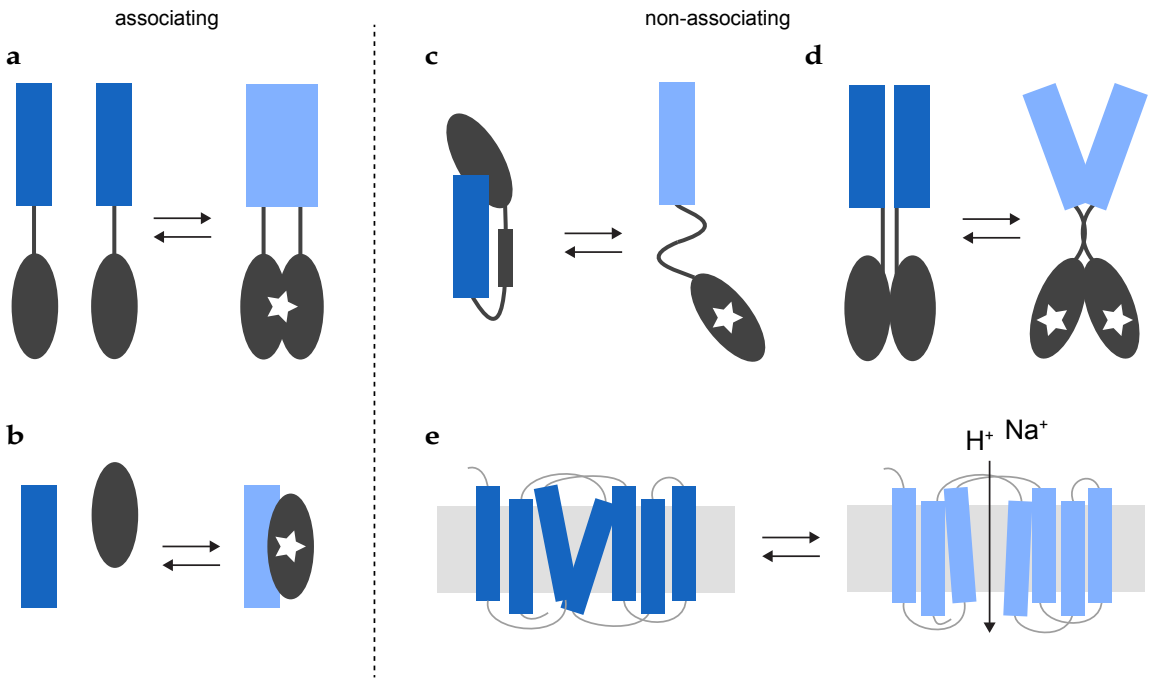

Figure 4: Allostery of photoreceptors (adapted from Ziegler and Möglich, 2015). Mechanisms of signal transduction in photoreceptors group into associating and non-associating types. The former change their oligomeric state in a light-dependent manner. Association can be either induced or repressed by light; can be between (**a**) alike (homo) or (**b**) different (hetero) partners; and can be of dimeric or higher stoichiometry. Non-associating photoreceptors comprise proteins undergoing light-regulated order-disorder transitions, e.g. of the Jα helix in *As*LOV2 (**c**), or tertiary/quaternary structural transitions (**d, e**). Dark and light colors of the sensor module indicate different photostates. Stars represent catalytic activity of the effector module.

which dissociate under far-red light[55,56]. In cryptochrome 2 (*At*Cry2) from the same plant blue light either elicits heterodimerization with CIB1[57] or oligomerization of *At*Cry2[58].

Non-associating photoreceptors comprise subtypes that undergo light-dependent order-disorder transitions (Figure 4c) or changes of their tertiary/quaternary structure (Figure 4d, e). The former class is represented by *As*LOV2[24] and the PYP (photoactive yellow protein)[59,60], where C- or N-terminal helices partially unfold in a blue-light dependent manner. The latter subtype includes microbial rhodopsins (Figure 4e), where light-induced rearrangement of the seven transmembrane helices induces ion transport across the membrane (e.g. channelrhodopsin)[2,61] or modulates interaction with associated G-proteins (e.g. human visual rhodopsin)[62]. The present study especially focuses on modular, homodimeric photoreceptor proteins, where helical linker regions relay light-induced structural changes in the photosensor module to alter the function of an effector module, like in the LOV protein YtvA from *Bacillus subtilis*[63] (Figure 4d).

The versatile light responses of LOV domains (3.1.1), e.g. in *As*LOV2 versus *Bs*YtvA-LOV, demonstrate that signaling mechanisms are not limited to a particular photochemistry. Depending on the respective protein architecture and effector module these mechanisms govern enzymatic activity, protein-protein interaction, DNA (deoxyribonucleic acid) binding and transcriptional control, ion flux, G-protein signaling or production of second messengers.

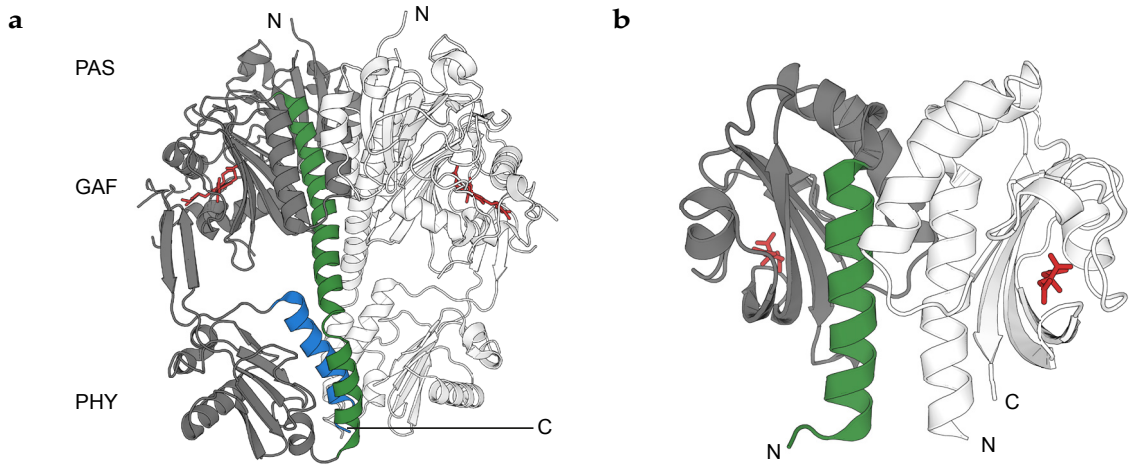

Figure 5: Recurring structural motifs in signal receptors. The (**a**) red-light receptor bacteriophytochrome from *Deinococcus radiodurans* (residue 4-503, PDB: 4O0P) and (**b**) the transmembrane citrate sensor CitA from *Klebsiella pneumoniae* (residue 4-132, PDB: 2J80) show parallel dimers with a central coiled coil.

3.2.1 Signal transduction in homodimeric signal receptors

Unlike homoassociating types, homodimeric photoreceptor proteins form dimers in a constitutive, light-independent manner. It should be noted that throughout this study the term 'homodimeric' describes modular, homodimeric receptors. Yet, non-modular proteins like channelrhodopsin probably form constitutive homodimers as well[64,65]. Structural data point toward prevalent parallel arrangements in prokaryotic types[10,39,66–68]; still anti-parallel orientations may occur in some cases[69,70]. A recurring feature among these proteins is a helical linker, the signaling helix, that thermodynamically couples sensor and effector modules[71]. In the dimer N- and C-terminal α helices form a central coiled-coil bundle that transmits structural changes from sensors to distal effector modules. In this way, multiple signals from consecutive sensors can be relayed and integrated to a single effector module by a continuous central helix spine[72,73]. Dimeric sensor modules that interact through their β sheets with a central helical spine, introduced for LOV and CBCR sensors, represent a common theme among diverse signal receptors. Apart from the superordinate PAS and GAF domains that also comprise sensors of various small metabolites, the PAS-GAF-PHY sensor module of phytochromes as well as the periplasmic sensor module of the transmembrane chemoreceptor CitA share the same overall architecture (Figure 5). Structural data on the mammalian transcription factor CLOCK:BMAL indicate that similar signaling principles might also occur in heterodimers and eukaryotes[74]. Membrane-spanning signaling helices and coiled coils are similarly abundant among transmembrane chemoreceptors and link extracellular or periplasmic sensors to intracellular effectors[75].

Structural rearrangements in the linker are still controversial and may involve translational, piston and rotational movements[76]. The signaling helix obviates direct contact between sensor and effector modules via specific, complementarity interfaces. Mutations in the sensor are therefore usually neutral allowing adaptation to a large variety of input signals[77] and a plethora of protein architectures featuring the same modules in various combinations. Structural changes

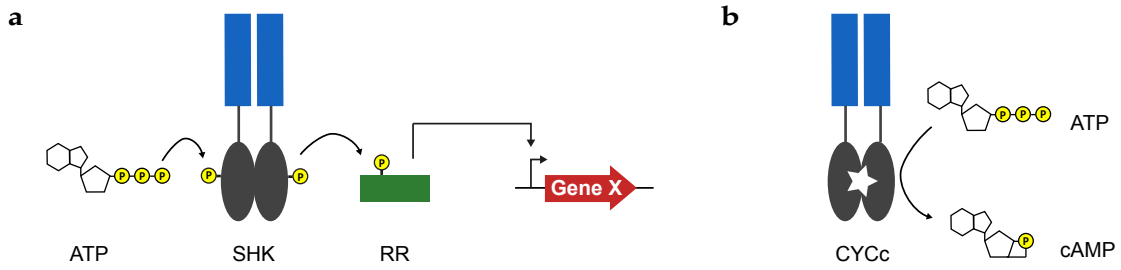

Figure 6: HisK and CYCc effector modules. (**a**) Homodimeric sensor histidine kinases (SHK) undergo ATP-dependent autophosphorylation and transfer the phosphate to a response regulator (RR), which in turn generates an output activity, e.g. changes expression of target gene. (**b**) Prokaryotic type III cyclases (CYCc) form catalytic sites at the dimer interface and convert ATP / GTP into the second messengers cAMP or cGTP.

propagated through the linker, affect the symmetry in the dimer and alter the effector activity through relative reorientation of the effector monomers or adjusting constraints within each monomer[67,78–80].

3.2.2 Effector modules of homodimeric signal receptors

Effector modules of homodimeric signal receptors are diverse[81,82] and include cyclases producing second messengers like cAMP / cGMP (cyclic adenosine or guanosine monophosphate) (CYCc) or c-di-GMP (GGDEF-EAL), phosphodiesterases (PDE) that degrade these cyclic nucleotides, histidine kinases (HisK), MCP (methyl-accepting chemotaxis protein) domains involved into chemotaxis and STAS (sulfate transporter / anti-sigma-factor antagonist) domains of unclear function. The present study especially deals with histidine kinases and adenylate / guanylate cyclases.

Together with a cognate response regulator (RR) protein, sensor-coupled histidine kinases (sensor histidine kinases, SHK) form two-component systems (TCS) – ubiquitous signal transduction systems in prokaryotes, which also occur in certain eukaryotes[83,84] (Figure 6a). The SHK usually catalyzes an ATP-dependent autophosphorylation on a conserved histidine residue. The phosphoryl group is then transferred to a conserved aspartate of the response regulator, which in turn elicits an output signal, often as changes in gene expression. In most cases, SHKs also catalyze the reverse reaction and act as phosphatases dephosphorylating the RR. Thereby, the input signal perceived by the sensor modulates the ratio between kinase and phosphatase activity. Hence, in this study the term 'kinase activity' denotes net excess of kinase activity of a protein compared to its phosphatase activity. SHK are usually highly specific for their cognate RR so that each two-component system is usually well insulated from other signal pathways in the cell. A common variation of the TCS is the phosphorelay where the SHK is a hybrid protein including a RR-like receptor domain that is phosphorylated before a histidine phosphotransferase transfers the phosphate to the RR. While SHKs are transmembrane, membrane-associated or cytosolic proteins and may feature diverse N-terminal sensor modules, the kinase effector itself comprises two highly conserved domains, the dimerization and histidine phosphotransfer (DHp) domain with the conserved histidine, and the catalytic and ATP binding (CA) domain.

Adenylate/guanylate cyclases[85] represent Type III of the adenylate cyclase family and occur in prokaryotes and eukaryotes[86]. In contrast to the other classes they form functional dimers with catalytic sites at the dimer interface. Thereby, prokaryotic cyclases usually form homodimers, whereas e.g. in mammalian Type III cyclases two different genes encode the monomers that assemble into pseudo-heterodimers with only one of them being catalytically active and the other exerting regulatory function. In combination with divalent cations as cofactors conserved residues in the catalytic site mediate the production of cAMP or cGMP from ATP or GTP (adenosine or guanosine triphosphate) (Figure 6b).

3.3 Photoreceptor engineering and optogenetics

In addition to their fundamental importance for organismal light adaption, sensory photoreceptors draw considerable attention due to their application as light-regulated switches in optogenetics[87]. Pioneered more than a decade ago[88], optogenetics combines optical and genetic techniques: heterologous expression of genetically encoded photoreceptors allows for control of cellular processes with the high spatiotemporal resolution of light application and minimal invasiveness. As the most prominent example, the light-gated cation channel channelrhodopsin-2 (ChR2) from *Chlamydomonas reinhardtii* is applied for depolarization of membrane potentials and induction of neuronal action potentials by light[89]. Therefore, ChR2 and homologues rhodopsins originating from diverse phototactic microorganisms represent a powerful tool to map neuronal circuits or control bioelectrical systems from neurons to heart tissue. Apart from the transmembrane rhodopsins, the cytosolic type III adenylate cyclase bPAC from *Beggiatoa sp*. facilitates blue-light induced production of intracellular cAMP[90] and enabled controlling sperm motility with light[91]. Complementary to their role as light-regulated actuators sensory photoreceptors, especially LOV proteins[92,93] and bacterial phytochromes[94,95], serve as fluorescent proteins (FP), whereas rhodopsins are employed as fluorescent reporters of membrane potential[96,97]. The optogenetic toolbox has been expanded by engineering sensory photoreceptors and rewiring functional modules into light-regulated chimeras[5,33].

A broad range of engineering approaches applies single point mutations (Figure 7a) to tweak spectral or kinetic properties of the intrinsic photochemistry. Well-known examples are non-sensory, fluorescent proteins derived from GFP (green-fluorescent protein), where extensive mutagenesis yielded variants of improved quantum yield, faster maturation and diverse spectral sensitivities. Great afford went into engineering channelrhodopsins to meet the challenges of their expansive application in neuroscience. Introducing mutations by semi-rational means generated variants with accelerated[98] or decelerated[99–101] channel kinetics, increased calcium conductance[102] and changed ion selectivity from cations to anions[103,104].
Another engineering approach is the generation of chimeras by exchanging homologous parts between proteins. Swapping transmembrane helices (Figure 7b) between different channelrhodopsins led e.g. to red-shifted absorption spectra[101,105]. Moreover, chimeras of light-sensitive rhodopsins and intracellular parts of homologous G-protein coupled receptors (GPCRs)

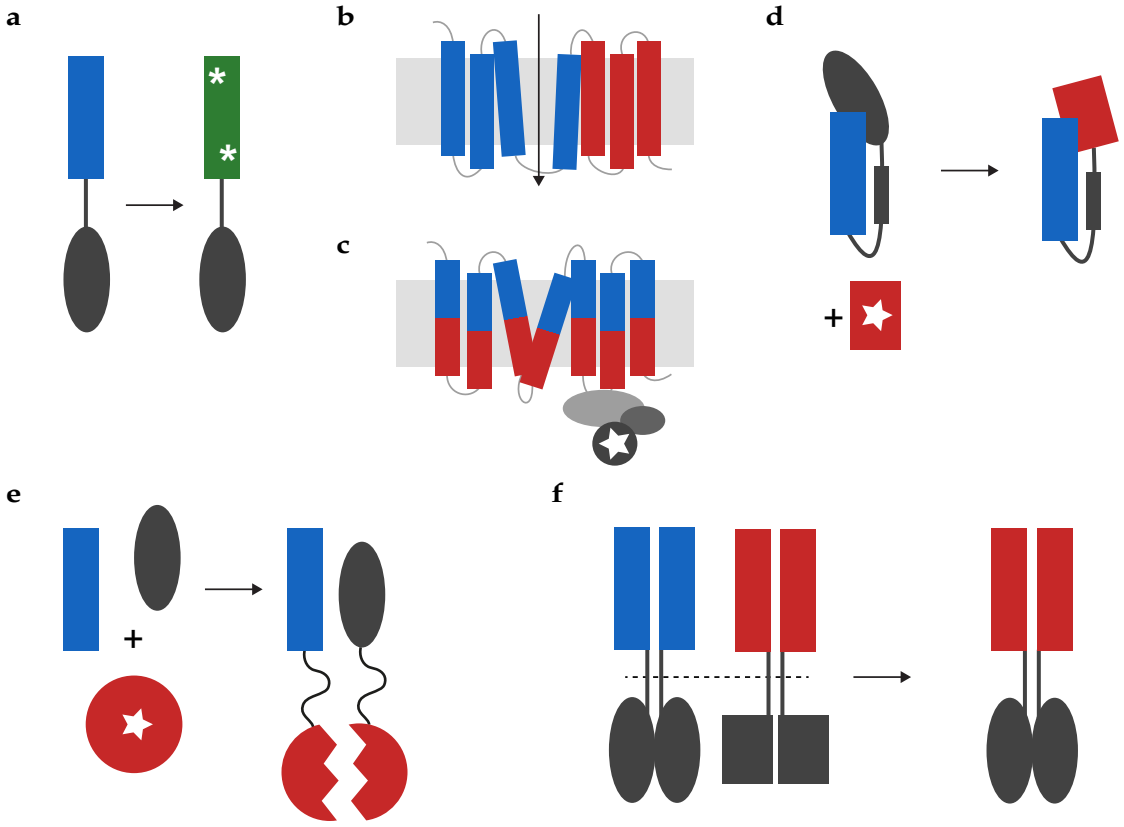

Figure 7: Photoreceptor engineering. (**a**) Site-directed mutation can optimize spectral sensitivities and kinetics of sensor modules. (**b**) Swapping transmembrane helices between different rhodopsins yield chimeras with desired traits. (**c**) Chimeras of light-sensitive and light-inert rhodopsins (opto-XRs) put G-protein signaling cascades under light control. (**d**) Light-induced order-disorder transitions, e.g. in *As*LOV2, facilitate reversible steric blocking of attached effector modules. (**e**) Associating photoreceptors allow restoring function of split effectors in a light-dependent manner. (**f**) Recombination of homologous sensor and effector modules in homodimeric receptors put target protein function under light control.

yielded so-called opto-XRs that facilitate light regulation of different G-protein signaling cascades[106,107] (Figure 7c).

Engineering modular, cytosolic photoreceptors for optogenetic purposes primarily relies on light-dependent steric blocking of effectors (Figure 7d) or split-protein approaches (Figure 7e). Various designs employ light-induced helix unfolding in AsLOV2 (Figure 4c) to reversibly mask signaling peptides or effector modules that directly control cellular processes[108–116] or recruit other effectors to do so[117–120]. Alternatively, *As*LOV2 was introduced into target proteins to disrupt their structure and function in a light-dependent manner[121].

Apart from *As*LOV2, most optogenetic tools use associating photoreceptors like VIVID[122–124], Phy/PIF[125–132], Cry2/CIB1[58,133–146], UVR8[129,147,148] or the non-sensory, fluorescent protein Dronpa[149] to recruit effectors to certain sites or restore split-protein architectures. Taken together, these approaches facilitate light regulation of gene expression[108,119,122,123,125,129,138,139,144–146,148,150–152], enzymatic activity[58,109,115,117,119,120,124,127,128,130,133–138,140–143,147,153], protein degradation[113,114], ion-channel activity[110,112,154], nuclear transport[116], intracellular transport[118] or apoptosis[111,123].

Moreover, along the lines of *what I cannot create, I do not understand*[155], engineering gives mechanistic information on the primary photochemistry and signal transduction within the

protein. Especially, the functional recombination of protein modules allows testing compatibility and shared signaling principles between the sensor and effector modules.

3.3.1 Engineered homodimeric photoreceptors

The architecture of homodimeric photoreceptors with structurally conserved modules recurring in various combinations points toward shared signaling principles between the modules and suggests their interchangeability (Figure 7f). Still, the linker between those modules, the signaling helix, remains the major design challenge. In natural and engineered associating photoreceptors the linker is usually unstructured and primarily fixes different protein modules in spatial proximity. Engineered *As*LOV2 photoswitches usually use the order-disorder transition of the original linker to make an effector peptide or module accessible in a light-dependent manner. In contrast, linkers of homodimeric photoreceptors form α helices that engage in specific interactions in the functional dimer and determine the relative position of sensor and effector modules. Moreover, they transmit a structural change from the sensor rearranging the effector modules in a defined manner to alter their activity level[71,75,78]. Consequently, the physical nature of the linker (topology, length, sequence, structure and dynamics) is decisive for functional coupling of sensor and effector modules and a successful design[5].

When recombining modules from different proteins, their original linkers have to be fused in a way that preserves signal transduction across the linker. Most engineering approaches exchange light-inert sensor modules from chemoreceptor proteins for photosensors. Following this strategy, Levskaya *et al.* replaced the histidine kinase of the phytochrome Cph1 from *Synechocystis sp.* with the homologous effector from the chemoreceptor EnvZ from *Escherichia coli* to yield Cph8

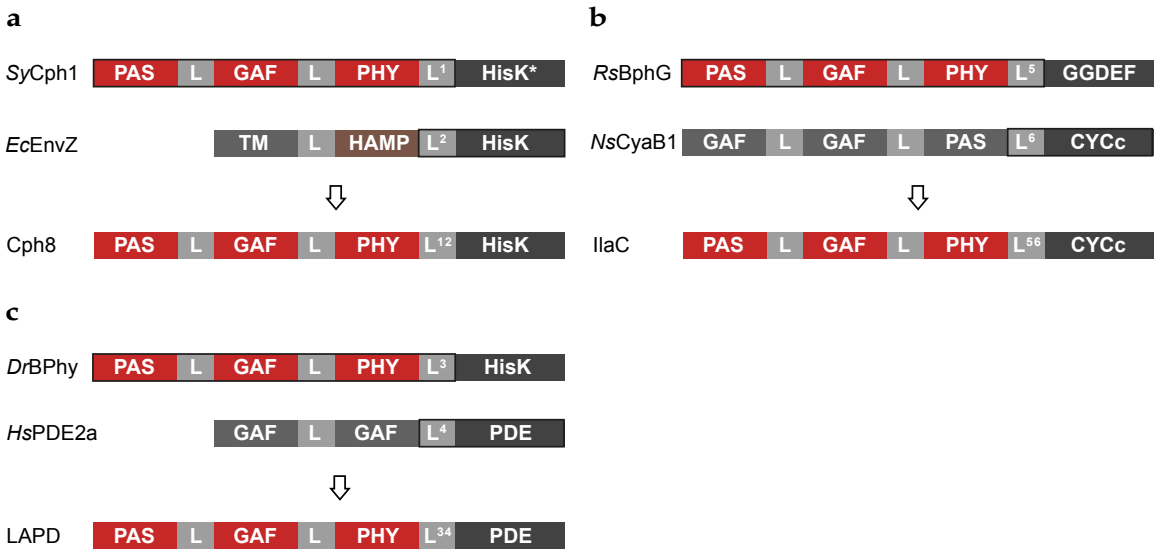

Figure 8: Domain architectures of engineered homdimeric photoreceptor chimeras. (**a**) The red-light-activated histidine kinase Cph8 (Levskaya *et al.* 2005). (**b**) The red-light-activated adenylate cyclase IlaC (Ryu *et al.* 2014). (**c**) The red-light-activated phosphodiesterase LAPD (Gasser *et al.* 2014). All chimeras employ either cyanobacterial or bacterial phytochromes. Species abbreviations: *Sy, Synechocystis sp.; Ec, Escherichia coli; Dr, Deinococcus radiodurans; Hs, Homo sapiens; Rs, Rhodobacter sphaeroides; Ns, Nostoc sp.*

(Figure 8a) and control gene expression in *E. coli* with red light[156]. Similarly, two other studies employed bacteriophytochrome sensor modules to create the red-light activated adenylate cyclase IlaC[157] (Figure 8b) and phosphodiesterase LAPD[158] (Figure 8c).

Möglich *et al.* replaced two PAS domains of the oxygen-regulated histidine kinase FixL from *Bradyrhizobium japonicum* by the structurally homologous YtvA-LOV domain (3.1.1) to yield the blue-light repressed histidine kinase YF1[78] (Figure 9a). The YF1 structure represents a rare full-length crystal structure of a constitutively homodimeric photoreceptor[70] and hence serves as a paradigm for investigating the signal transduction between sensor and effector modules. The structure comprises the N-terminal YtvA-LOV domain connected to the histidine kinase (dimerization and histidine phosphotransfer (DHp) + catalytic (CA) domain) via the helical Jα linker[67] (Figure 9b). Within the functional YF1 homodimer, the two Jα helices assemble in a coiled coil. The significant kink in the structure supposedly arises from crystal-packing forces[67]. Experiments *in vitro* as well as *in vivo* showed net kinase activity of YF1 in the dark, which switches to a net phosphatase activity upon blue-light illumination[78]. Spectroscopic and biochemical data led to a minimal model where kinase activity is only observed when both LOV domains are in the dark state[78]. On the structural level blue-light absorption induces conformational changes within the LOV domain that are suggested to elicit rotation of the Jα helices, which in turn rearranges the kinase modules and modulates their activity. A similar model was proposed for the antiparallel four-helix bundle of HAMP (histidine kinase, adenylate

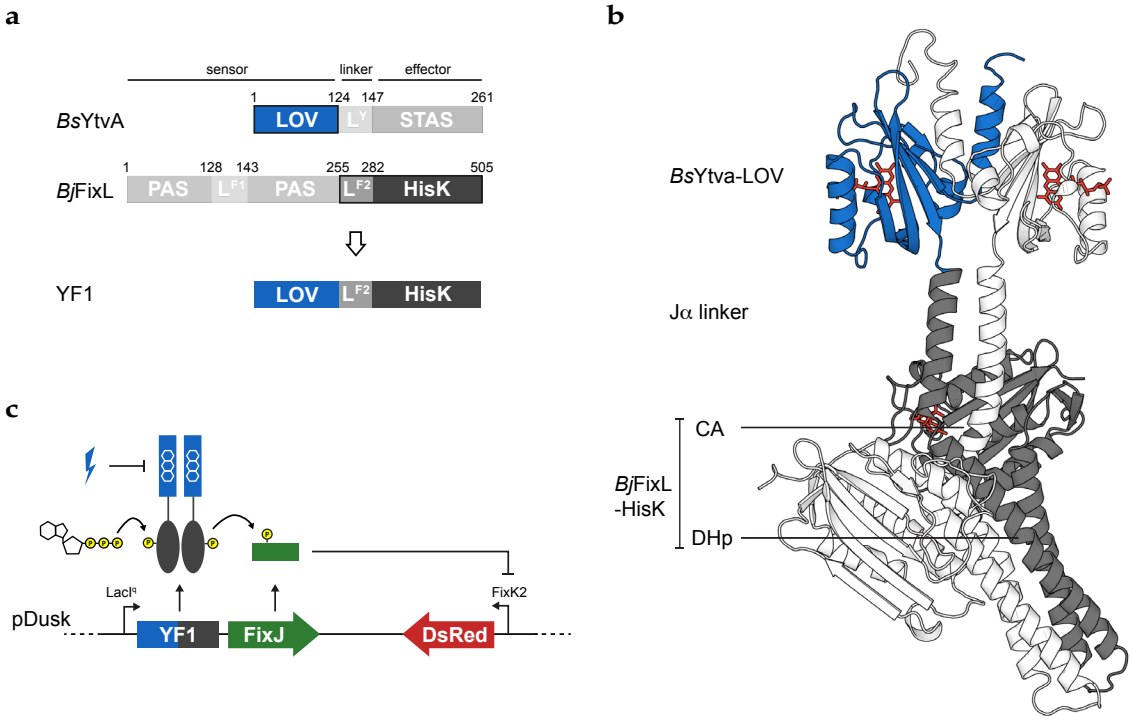

Figure 9: The engineered photoreceptor YF1. (**a**) YF1 comprises the *Bs*YtvA-LOV domain and the *Bj*FixL-histidine kinase including its linker. (**b**) The crystal structure of YF1 (PDB: 4GCZ) displays the central coiled-coil spine linking LOV-sensor and HisK-effector domains in the homodimer. (**c**) In the pDusk reporter plasmid kinase activity of YF1 phosphorylates (yellow circles) its response regulator FixJ, which then triggers expression of red-fluorescent DsRed from the FixK2 promoter. Kinase activity of YF1 is largely diminished by blue-light illumination.

cyclase, methyl accepting protein and phosphatases) domains, which primarily occur in transmembrane chemoreceptors[75].
YF1 serves as basis for the plasmids pDusk and pDawn that permit regulating gene expression in *Escherichia coli* in a light-dependent manner[159]. The pDusk plasmid comprises YF1 and the cognate response regulator FixJ that are constitutively expressed as bicistronic operon (Figure 9c). In the absence of blue light, YF1 acts as a histidine kinase and phosphorylates FixJ, which then induces gene expression from the FixK2 promoter that is situated immediately upstream of a multiple-cloning site (MCS). Blue-light application inhibits the kinase function of YF1 and thereby represses expression from the FixK2 promoter. Introducing a red-fluorescent reporter into the MCS, pDusk enabled functional testing of YF1 mutants[67,160,161].

3.4 Aim of the project

Modules facilitating light perception and exertion of biological activity recur in sensory photoreceptors in various combinations, which points toward overarching principles of signal transduction. Engineered chimeric photoreceptor proteins support the notion that distinct photosensor and effector modules can be functionally rewired. Furthermore, engineering extends the optogenetic toolbox by putting target cellular signaling pathways under light control. A key challenge of the design process is the linker region that physically and functionally couples sensor and effector modules. Recombining different modules requires fusion of the original linkers from the parent proteins. While linkers of associating photoreceptors are usually unstructured, the helical linkers of homodimeric photoreceptors form a central-coil spine in the functional dimer (3.2.1). Beyond bringing sensor and effector modules into spatial proximity, they transmit the signal as a structural change. Therefore, their physical properties crucially determine sensor-effector communication and functionality of engineered chimeras. Indeed, the knowledge about the mechanisms of signal transduction through the linker and their conservation among different module classes is scarce. Hence, the design process is often unpredictable and time-consuming.
The present study aims to explore efficient strategies for rewiring modules of homodimeric receptor proteins. First of all, photosensor modules are functionally connected to new effectors based on sequence similarity and structural cues. I start out from the FixL-histidine kinase where previous extensive functional and structural studies should greatly benefit the recombination with novel sensor modules. The focus lies on the spectrally diverse and photochromic cyanobacteriochrome photosensors (3.1.2). Both functional properties as well as their small size comprising a single GAF domain make them highly attractive for optogenetic applications. While spectral diversity could lead to chimeric receptors with orthogonal light sensitivities, photochromic photocycles allow induction and inhibition of protein activity by the application of two different wavelengths. Structural data indicate that CBCRs form functional homodimers like other GAF sensors. Moreover, they occur in combination with histidine kinases in natural proteins indicating their general compatibility with this effector type and making CBCR-FixL fusions a natural target (5.1). Thereby, the pDusk plasmid (Figure 9c) offers a reporter system for testing numerous fusion constructs with little effort. Going step-by-step from conservative to

more complex designs, also other effector modules should be rewired to CBCR sensors. Adenylate/guanylate cyclases are suitable targets, because they work as homodimers and generate the important intracellular messengers cAMP or cGMP (5.2).

While these studies follow a rational design approach, a second part of this work focuses on developing a complementary brute-force strategy. Instead of structure- and sequence-guided planning of fusion constructs, the novel approach should generate fusion libraries of a sensor-effector pair with random combinations of their original linkers. Bacterial reporter assays, like pDusk for FixL fusions, will then facilitate quickly testing large numbers of generated fusion variants. Beyond accelerating the design-test cycle for creating chimeric photoreceptors, functional chimeras will map out the relation of linker properties and protein function.

Hence, the project comprises two complementary and interdependent parts: On the one hand, the characterization of functional chimeric photoreceptors contributes to an improved mechanistic understanding of signal perception and signal transduction. On the other hand these studies will contribute strategies for further engineering photoreceptors applicable in optogenetics.

4 Materials and Methods

4.1 Biological and chemical materials

Unless stated otherwise all enzymes were purchased from Thermo Fisher Scientific Inc. (St. Leon-Rot, Germany). All reagents were purchased from Carl Roth GmbH & Co. KG (Karlsruhe, Germany) or VWR International GmbH (Dresden, Germany) and were of grade „Molecular Biology“ or higher. All *Escherichia coli* strains and plasmids used in this study are listed in Table 2 and Table 3. Please note that in this study ‘pDusk’ denotes the original pDusk plasmid[159] with an additional Myc-tag at the end of the kinase gene and DsRed-Express2 cloned into the multiple-cloning site (pDuskMM-DsR)[67] (Figure 9c).

Table 2: *Escherichia coli* strains applied in protein expression and reporter assays.

Strain	Genotype
DH10b	F- mcrA Δ(mrr-hsdRMS-mcrBC) Δ80lacZΔM15 ΔlacX74 recA1 endA1 araΔ139 Δ(ara, leu)7697 galU galK Δ-rpsL (StrR) nupG
CmpX13	F – ompT hsdSB (rB- mB-) gal dcm (DE3) manX::ribM
CmpX13ΔcyaA	F – ompT hsdSB (rB- mB-) gal dcm (DE3) manX::ribM Δcya
BTH101	F- cya-99 araD139 galE15 galK16 rpsL1 (StrR) hsdR2 mcrA1 mcrB1
BL21 (DE3)	F- ompT hsdSB (r-Bm-B) gal dcm (DE3)
C41 (DE3)	F – ompT hsdSB (rB- mB-) gal dcm (DE3)

Table 3: Plasmids used in protein expression and reporter assays. Respective selection markers mediate resistance against Kanamycin (Kan), Ampicillin (Amp), Chloramphenicol (Cam) or Streptomycin (Strep).

Plasmid	Origin	Selection marker	Application / Manufacturer or parent plasmid
pET28c	pBR322	Kan	Expression vector / Novagen
pASK	pBR322	Amp	Expression vector / IBA
pACYC177	p15a	Amp	Cloning vector / NEB
pACYC184	p15a	Cam	Cloning vector / NEB
pCDF-Duet	CloDF13	Strep	Expression vector / Novagen
pBADM30	ColE1	Amp	Expression vector / EMBL Heidelberg
pDuskMM-DsR[54,130]	ColE1	Kan	HisK test vector / pET28c
pCyclR	ColE1	Kan	CYCc test vector / pET28c
pKT271	p15a	Cam	PCB synthesis / pACYC184
pChromCDF	CloDF13	Strep	PCB synthesis / pCDF-Duet1
p171	ColE1	Amp	PCB synthesis

4.2 Molecular biology

4.2.1 Restriction cloning

Target DNA inserts were produced by digestion of the parent plasmid or by PCR (polymerase chain reaction) amplification (Table 4). For cloning of constructs, plasmid DNA and PCR products were digested for 10′ to 30′ with the respective restriction enzymes in the appropriate digestion buffer (Fast Digest, Thermo Fisher Scientific, St. Leon-Rot, Germany) (Table 5a). Plasmids were dephosphorylated by incubation with alkaline phosphatase (Thermo Fisher Scientific) for 15′ at 37°C (Table 5a). The digested plasmids, inserts and PCR products were checked via agarose-gel electrophoresis and afterwards purified by gel extraction or PCR-clean up (NucleoSpin Extract II, Machery-Nagel GmbH & Co. KG, Düren, Germany). DNA ligation of plasmids and inserts was carried out at a molar ratio of 1:6 (total volume 8 µl) using the T4 DNA Ligase in the corresponding buffer (Thermo Fisher Scientific) and incubation for at least 1 h at 22°C (Table 5b). The complete reaction mix was transformed into chemically competent *E. coli* DH10b via heat shock. After incubation for 1 h at 37°C cells were streaked onto an LB (Lysogeny broth) agar plate

Table 4: Standard PCR protocol. (**a**) Thermo cycler protocol and (**b**) PCR-reaction mixture.

a

Temperature	Time	
98°C	30 s	
98°C	30 s	25 cycles
Tm-5	30 s	25 cycles
72°C	20s/kb	25 cycles
72°C	10 min	
10°C	∞	

b

Amount	Substance
10-50 ng	Template DNA
0.2 mM each	dNTPs
10 µl	5x HF Buffer
0.2 µM	Fwd Primer
0.2 µM	Rev Primer
1 µl	Phusion Polymerase
to 50 µl	ddH_20

Table 5: DNA digestion and ligation. (**a**) Composition of reaction mixture for DNA digestion. The phosphatase is only added to plasmid DNA. (**b**) Reaction mixture for DNA ligation.

Amount	Substance
3.4 µl	FD Buffer
1.5 µl	Restriction Enzyme I
1.5 µl	Restriction Enzyme II
30 µl	DNA
1.5 µl	FastAP

Amount	Substance
40-100 ng	Plasmid DNA
variable	Insert DNA
1 µl	Ligase buffer
1 µl	T4 DNA Ligase
to 10 µl	ddH_20

(1% agar) supplemented with the appropriate antibiotics (standard concentration were: 50 μg ml^{-1} kanamycin and streptomycin and 30 μg ml^{-1} for ampicillin and chloramphenicol) and grown at 37°C for at least 12 h. 5 ml cultures were inoculated with single colonies and grown for at least 16-21 h at 37°C. Afterwards plasmid DNA was purified (NucleoSpin Plasmid MiniPrep kit, Machery-Nagel) and sequences were checked by both analytical digest and DNA sequencing carried out by GATC Biotech AG (Constance, Germany) or LGC Genomics (Berlin, Germany). Correct plasmids were transformed into the respective expression or assay strain and stored as glycerol stocks at -80°C.

4.2.2 Gibson cloning

Fragments of target genes and plasmid backbone were amplified via PCR as described above (Table 4). Thereby, oligonucleotide primers applied for amplification of the gene insert featured overlaps of around 20 bases to the insertion site in the plasmid. Both fragments were purified via gel extraction and fused via Gibson assembly[162]. A reaction master mix (Table 6b) was prepared from 5x reaction buffer (Table 6a) and T5 exonuclease (Epicentre), Phusion polymerase (NEB, New England Biolabs GmbH, Frankfurt am Main, Germany) and Taq ligase (NEB) and stored as aliquots of 15 μl at -20°C. To a thawed aliquot insert and plasmid DNA was added in equimolar amounts (or at a 5-fold molar excess when the insert was considerably smaller) to a total volume of 20 μl. The reaction mix was incubated at 50°C for 1 h followed by 15 min at room temperature and 2 μl were then transformed into chemically competent *E. coli* DH10b.

Table 6: Gibson cloning. (**a**) Composition of 5x reaction buffer and (**b**) reaction master mix.

a

Amount	Substance
500 mM	Tris-HCl pH 7.5
50 mM	$MgCl_2$
50 mM	DTT
25 %	PEG-8000
1 mM each	dNTP
5 mM	NAD

b

Amount	Substance
4 μl	5x reaction buffer
0.64 μl	T5 exonuclease (10 U μl^{-1})
20 μl	Phusion polymerase (2 U μl^{-1})
160 μl	Taq ligase (40 U μl^{-1})
To 1.2 ml	ddH_2O

4.3 Bacterial histidine-kinase and adenylate-cyclase assays

Chimeric proteins were functionally tested using fluorescence-reporter assays in *Escherichia coli*. Histidine kinases were cloned into pDusk[67,159] (Figure 9c), where catalytic activity induces expression of red-fluorescent DsRed-Express2[163]. Adenylate cyclases were cloned into pBAD and

co-transformed with the pCyclR plasmid (5.2.1), which leads to DsRed expression upon cAMP production by the cyclase. Unless stated otherwise, cells were grown in Lysogeny broth (LB) supplemented with respective antibiotics. Standard concentrations were 50 μg ml^{-1} for kanamycin and streptomycin and 30 μg ml^{-1} for ampicillin. Cells harboring the construct of interest were grown in 5 ml LB supplemented with the appropriate antibiotics for 18 h at 37°C. Aliquots of 6 μl were used to inoculate 96-deep-well plates containing 600 μl LB per well and eventual inducing agents. The plates were sealed with gas-permeable film and incubated for 22 h at 37°C and 800 rpm in the dark or under blue light (470 nm, 40 μW cm^{-2}) (Figure 10b). Light was applied via custom-built arrays of 10 x 8 LEDs of 470 ± 10 nm, 525 nm ± 10 nm or 650 nm ± 10 nm (Winger Electronics GmbH & Co. KG, Dessau, Germany). Light intensities were measured using a power meter (model 842-PE, Newport, Darmstadt, Germany) and a silicon photodetector (model 918D-UV-OD3, Newport). Afterwards, the cultures were diluted 20-fold in the case of kinases and 26-fold for cyclases with distilled water in black-walled 96-well μClear plates (Greiner BioOne, Frickenhausen, Germany). Absorbance at 600 nm and DsRed fluorescence (excitation 554 ± 9 nm, emission 591 ± 20 nm) was measured using a Tecan Infinite M200 PRO plate reader (Tecan Group Ltd., Männedorf, Switzerland). Data were normalized to the fluorescence per A_{600} of YF1 in pDusk under dark conditions and represent average values of two biological replicates ± standard deviation. Please note that for the sake of simplicity throughout this study I speak of histidine kinase or adenylate cyclase activity measured by the described assays. Nevertheless, measurements of DsRed fluorescence output of both test systems give indirect information on the catalytic activities of the tested proteins. Furthermore, histidine kinases usually exert both phosphorylation and dephosphorylation of their cognate RR, so that 'kinase activity' always means a net excess of kinase activity compared to the phosphatase activity of a protein.

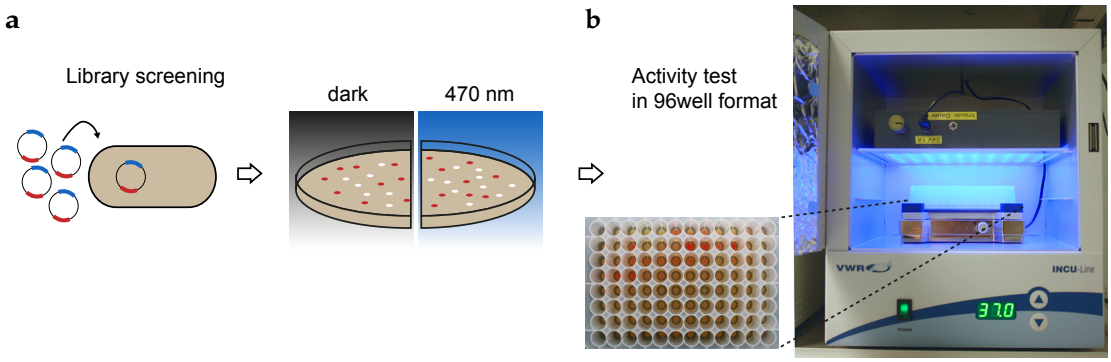

Figure 10: Functional testing with bacterial assays. (**a**) Fusion libraries are transformed into *E. coli* and grown on agar plates under desired test conditions, here dark and blue light. Bacterial reporter assays involve production of red-fluorescent DsRed in response to histidine-kinase or adenylate-cyclase activity. (**b**) Clones harboring active fusion proteins are selected based on their red color and sorted into 96-well plates containing LB. After overnight growth in the dark or under blue light, DsRed fluorescence is measured and compared under both conditions.

4.4 Generation of linker libraries

The starting construct YF was created by fusing the genes encoding residues 1-147 of *Bacillus subtilis* YtvA and residues 255-505 of *Bradyrhizobium japonicum* FixL with an *NheI* site in between via overlap-extension PCR and cloned into pDusk (GenBank JN579120). A frame shift at the fusion site was introduced to yield a strictly inactive starting construct. The starting constructs YCy and YP comprising the residues 1-147 of YtvA and 564-859 of CyaB1 from *Anabaena sp.* or 95-305 of bPAC from *Beggiatoa sp.* were cloned into pBAD.

4.4.1 ITCHY libraries

ITCHY libraries were generated using a modified version of the published protocol[164]. Applied enzymes mostly work in similar buffers to streamline the workflow and eliminate additional purification steps. The YF-pDusk template was digested with FastDigest *NheI* and the linearized plasmid was isolated via gel extraction. The open ends were incrementally digested using 100 U exonuclease III per μg DNA at 4°C. Aliquots were taken at 10, 30, 60, 120, 300 s and the reaction immediately stopped with S1 nuclease reaction mix (S1 nuclease in reaction buffer). The samples were kept on ice until the last aliquot was taken and then moved to room temperature for 30 min to initiate blunt-end generation by S1 nuclease. All samples were pooled and purified via PCR-clean up. By using a mixture of T4 DNA Polymerase, Klenow fragment and T4 Polynucleotide kinase (End-repair kit) blunt-end generation was further enhanced and the open plasmid ends phosphorylated. After another PCR-clean up the plasmids were ligated with high-concentration T4 DNA Ligase (30 U μl^{-1}). Remaining template plasmids were removed by *NheI* digestion followed by dephosphoralytion with alkaline phosphatase (FastAP) and the ligated plasmids were again purified via PCR-clean up.

4.4.2 PATCHY libraries

PATCHY libraries were generated using forward and reverse sets of staggered primers annealing in the sensor or effector linker regions of the respective starting construct (YF-, YCy- or YP-pDusk). Each set contains primers starting at every triplet of the respective linker region. The primers were designed with a custom Python script (in collaboration with Dr. Florian Richter, HU Berlin), which can be downloaded at https://github.com/vrylr/PATCHY.git. Briefly, the script designs primers for a set of start positions in the template sequence, which have similar annealing temperatures (here set to 60°C) and certain maximal and minimal lengths. With one forward primer for each triplet in the effector linker (27 residues) and one reverse primer for each triplet in the sensor linker (23 residues), 52 oligonucleotide primers generate 672 (including constructs where either one of the linkers or both are completely removed: (23+1) · (27+1)) YtvA-FixL fusion variants. In general for two linkers with n and m residues, $n+m+2$ primers generate $(n+1) \cdot (m+1)$ fusion variants. Forward and reverse primers were pooled and applied at a final concentration of 0.5 μM in a PCR (Table 4) with the YF-pDusk template. The PCR product was purified via PCR-clean up and supplemented with T4 DNA ligase buffer. For phosphorylation and ligation the linear plasmids were incubated with T4 Polynucleotide kinase for 30 min at 37°C, supplemented

with PEG4000 and T4 DNA Ligase and incubated for another hour at room temperature. Remaining template plasmids were removed by addition of *NheI* and FastAP. The applied enzymes all have full activity in T4 DNA ligase buffer, so that no purification steps are needed between the different reactions. Purification via PCR-clean up after the complete procedure is recommended to increase DNA stability upon storage.
Naïve PATCHY libraries were analyzed by next-generation sequencing (NGS, GATC Biotech AG). Samples were prepared in collaboration with Charlotte H. Schumacher (HU Berlin). Plasmid libraries were randomly fragmented by ultrasound and sequenced on an Illumina platform that yields paired-end reads of 125 base pairs length. In total, about $2 \cdot 5{,}500 = 11{,}000$ reads were obtained that correspond to in-frame fusions of *B. subtilis* YtvA and *B. japonicum* FixL. Given a theoretical library size of 672 linker variants, this corresponds to about 8-fold library coverage. Prof. Andreas Möglich evaluated sequence data using Python and Numpy (http://www.numpy.org/), and graphs were prepared with Matplotlib (http://matplotlib.org/).

4.4.3 Library screening

ITCHY or PATCHY libraries of kinase fusions (YtvA-FixL) were transformed into *E. coli* DH10b and kinase activity was evaluated using pDusk as described before[159,160] (Figure 10). PATCHY libraries of cyclase fusions (YtvA-CyaB1 and YtvA-bPAC) were first transformed into DH10b. After overnight growth on agar plates approximately 10^5 colonies were pooled and plasmid DNA was isolated. These libraries were then transformed into CmpX13ΔcyaA and cyclase activity was evaluated using the co-transformed pCyclR reporter plasmid (5.2.1). The procedure of library screening was similar for both, kinase and cyclase libraries. First, the transformed cells were grown on agar plates in the dark or under blue light (470 nm, 40 μW cm^{-2}). Clones harboring active kinase or cyclase fusions were then selected based on their DsRed fluorescence (excitation 470 nm) using a 590 nm high-pass filter and sorted into 96-deep-well plates containing 600 μl LB/Kan for kinases or LB/Amp+Kan for cyclases and grown for 18 h at 37°C and 800 rpm. Aliquots of the cultures were mixed with glycerol for long-term storage at -80°C. Aliquots of 6 μl from each well were used to inoculate two replica plates for light/dark testing as described before. Light-switchable constructs were identified by comparing DsRed levels under dark and light conditions and submitted to DNA sequencing (GATC Biotech). Sequencing data was analyzed with custom Python scripts (download at https://github.com/vrylr/Seq_analysis.git) and ClustalX. All verified clones were again tested as duplicates in the 96-well setup so that the data represent average values of two biological replicates ± standard deviation.

4.5 Protein expression and purification

Cells (BL21, CmpX13 or C41) harboring the respective expression and accessory plasmids were grown in 800 ml LB at 37°C and 225 rpm to an OD_{600} of 0.4 in an incubator shaker (225 rpm, Innova 44R, New Brunswick Scientific, Edison, NJ). Depending on the construct, protein expression from pET28c and the chromophore plasmids pKT271 and p171 was induced by addition of 1 mM IPTG. Incubation continued over night at 18°C. Cells were harvested by

centrifugation for 5 min at 7000 rpm using a JA10 rotor (Beckman Coulter GmbH, Krefeld, Germany). The pellet was resuspended with 2 x 7 ml buffer A (50 mM Tris/HCl (pH 8.0), 20 mM NaCl, 20 mM imidazole) supplemented with protease inhibitor mix (Roche protease inhibitor, 1 pill per 40 ml; Roche Diagnostics Germany, Mannheim, Germany) and cells were lysed by sonication. Insoluble cell fragments were removed by centrifugation for 30 min at 18000 rpm using a JA25.5 rotor (Beckman Coulter GmbH) and the lysate was purified by metal ion affinity chromatography using 1 ml of HisPur Cobalt resin (Thermo Scientific). After loading the supernatant onto the column (equilibrated with buffer A) it was washed with 40 ml buffer A, 20 ml buffer C (50 mM Tris/HCl (pH 8.0), 1 M NaCl) and 20 ml of buffer B (50 mM Tris/HCl (pH 8.0), 50 mM NaCl). The hexahistidine-tagged proteins were eluted with 1 ml aliquots of buffer D (50 mM Tris/HCl (pH 8.0), 20 mM NaCl, 200 mM imidazole) and samples were analyzed via SDS-PAGE (12% agarose gel supplemented with 1 mM zinc acetate). Incorporation of bilin chromophores could be confirmed via fluorescence imaging of the zinc-bilin complex. Fractions of highest protein amount and purity were pooled and dialyzed against 10 mM Tris/HCl (pH 8.0), 10 mM NaCl. The protein concentration was evaluated via absorption using extinction coefficients at 650 nm for *Np*R5113_GAF2 (ε = 98000 cm^{-1} M^{-1})[165] and Lambert-Beer's law.

UV-vis spectra were recorded with an Agilent 8453 spectrophotometer (PDA, Agilent Technologies, Waldbronn). Samples were illuminated with LED light of 525 nm or 650 nm. The spectra were corrected for the baseline signal between 800-900 nm.

4.6 Linker-length analysis

For a linker-length analysis of natural receptor proteins identifiers of proteins containing PAS (PF00989, PF13426, PF08447, PF08448, PF13188, PF08348, PF12860, PF13596, PF16527, PF08446, PF14598, PF16736, PF07310, PF08670), GAF (PF01590, PF13185, PF13492), BLUF (PF04940) or PHY (PF00360) sensor modules or HisK (PF00512, PF07536, PF07568 and PF07730), GGDEF (PF00990), CYCc (PF00211) or PDE (PF00233) effector modules were retrieved from the Pfam database[82]. Due to the large dataset, PAS-GGDEF analysis focused on a subset of 2000 sequences. Respective outcomes however resembled previous data[166]. Custom Perl scripts supported all following steps[78] (provided by Prof. Dr. Andreas Möglich, Universität Bayreuth). Sequence data of proteins that appeared in target sensor and effector lists, e.g. PAS and HisK, were retrieved from the Uniprot database[167]. Resulting sequences were filtered for protein architectures where the sensor resides N-terminally to the effector in a maximal distance of 100 amino acids. This threshold should eliminate proteins featuring additional domains between the target sensor and effector modules. A disadvantage is that it also eliminates potential proteins with exceptionally long linkers. The protein sequences were cropped leaving the linker and 15-20 neighboring residues, aligned using ClustalX[168] and the linker lengths counted between two sequence motifs of choice. Prof. Dr. Andreas Möglich performed the analysis of PAS-HisK proteins. Linker analysis of GAF-HisK proteins for planning CBCR-HisK fusions in section 5.1.1 were restricted to GAF subclass PF01590 and HisK PF00512 that comprise CBCRs and the target *Bj*FixL histidine kinase.

5 Results

5.1 Rational design of CBCR-HisK fusions

5.1.1 Direct CBCR-FixL fusions

The diversity of cyanobacteriochromes poses the question of which are most suitable for design purposes. While the photosensor modules of all CBCRs belong to the GAF-domain family and possess the typical GAF fold (3.1.2), they largely differ regarding most other properties. Here, the relevant selection criteria were spectral sensitivity, chromophore type and good heterologous expression in *E. coli*. To fully utilize their bistability, I focused on sensors with spectrally well-separated absorption states, like the red-green cyanobacteriochromes, which usually show two states with peak absorptions more than 100 nm apart[165]. Moreover, a red-light sensitive photoswitch would be orthogonal to most, often blue-light sensitive, photoreceptors applied in optogenetics. While CBCRs use PCB or PVB chromophores, the latter has the drawback that it needs additional maturation inside the protein leading to heterogeneous chromophore populations[45]. Therefore, the choice was narrowed down to PCB-binding CBCRs. Finally, the proteins with high expression in *E. coli* (personal communication with Nathan Rockwell, UC Davis) were selected from the remaining candidates leading to two PCB-binding, red-green sensitive domains, GAF2 of R5113 and GAF6 of F2164 both from the cyanobacterium *Nostoc punctiforme*[165].

In order to get a better understanding of fusing CBCR sensors to the FixL histidine kinase the linker regions of natural GAF-HisK proteins were analyzed. Using the Pfam database[82] all proteins harboring either a GAF domain or a histidine kinase were identified and sequences of proteins that occur in both data sets were retrieved from the Uniprot database[167]. I restricted the search to the subclasses of GAF domains and histidine-kinases that comprise CBCRs and FixL

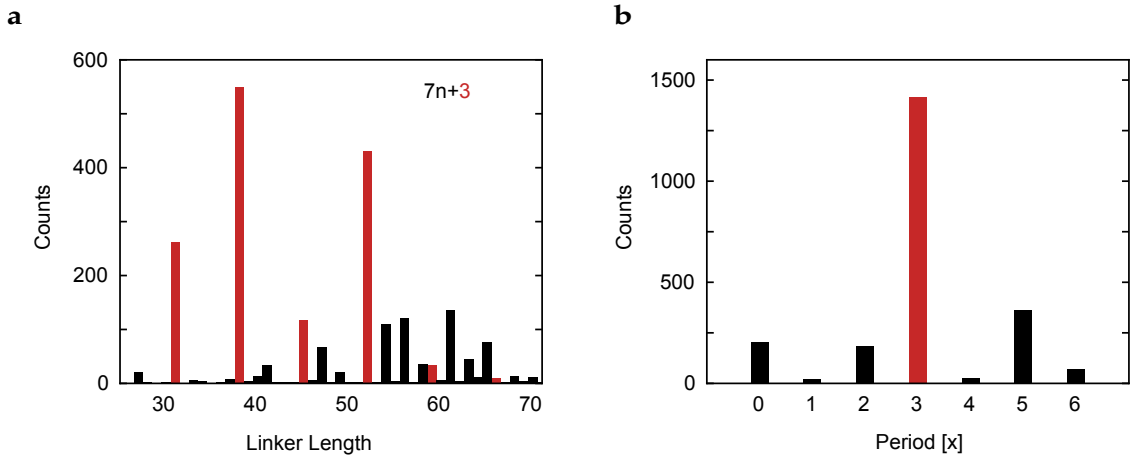

Figure 11: Linker analysis of natural GAF-HisK proteins. (**a**) Alignment of natural proteins containing GAF domains (Pfam: PF00512) situated N-terminally of a histidine kinase (Pfam: PF01590) display preferred linker lengths. (**b**) Linker lengths distribute according to a heptad periodicity of 7n+3 residues (highlighted in red)

(PF01590 and PF00512). Defining criteria, like a maximal distance and a certain order of the modules, narrowed data down to GAF-HisK proteins that contain an N-terminal GAF sensor directly linked to a C-terminal histidine kinase. The resulting alignment shows that linkers between the two modules usually consist of 7n+3 residues, e.g. 38, 45 and 52 *et cetera* (Figure 11). Similar heptad periodicity was previously reported for PAS-histidine kinases[78] and might imply a conserved mechanism of signal transduction between sensor and effector modules among these histidine kinases. Consequently, in the CBCR-HisK fusions G2F (*Np*R5113 206-365 + *Bj*FixL 260-505) and G6F (*Np*F2164 1037-1200 – *Bj*FixL 262-505) the parental linkers were combined into linkers of 38 residues length, which represents the predominant linker length among natural GAF-HisKs and closely resembles the linker length in the parent FixL protein (Figure 12a, b, Table 7).

Heterologous expression of functional CBCRs needs exogenous supply of the PCB chromophore.

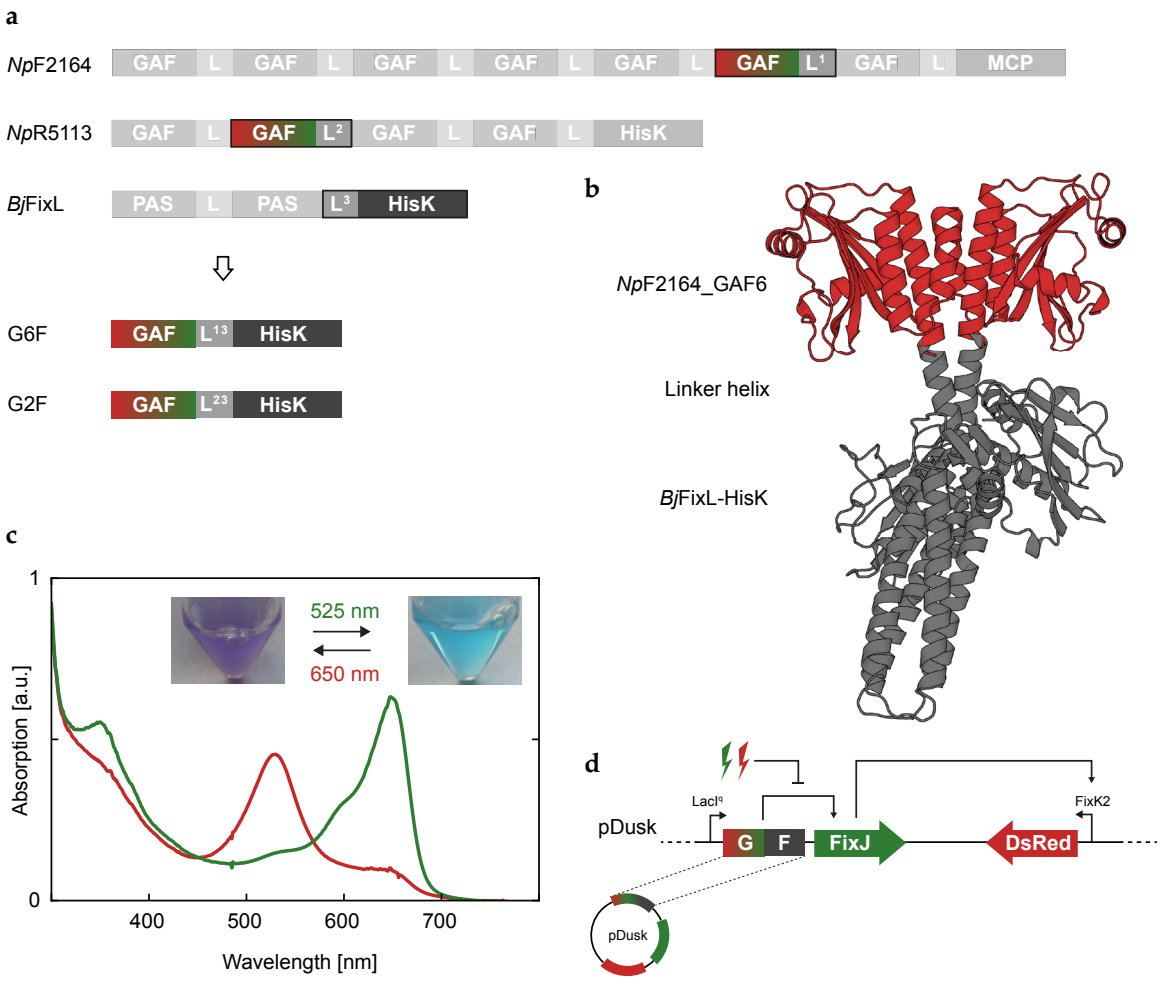

Figure 12: Direct CBCR-FixL fusions. (**a**) The red-green sensitive GAF6 of *Np*F2164 and GAF2 from *Np*R5113 were rewired to the *Bj*FixL-histidine kinase by fusing the parental linkers, yielding the chimeras G6F and G2F. (**b**) Structural alignment of *Np*F2164_GAF6 (residues 1038-1203, homology model based on 3W2Z (swissmodel.expasy.org) and *Bj*FixL. (**c**) Absorption spectra of purified G2F exhibit the equilibrium between green- and red-absorbing states that can be shifted to either side by application of red and green light. A visible color change accompanies the transition (small insets). (**d**) CBCR-FixL fusions are tested for kinase activity using the pDusk reporter.

The accessory proteins hemeoxygenase 1 (HO-1) and bilinreductase (PcyA) can be provided on an additional plasmid to convert endogenous heme into PCB in *E. coli*[169,170]. Two different plasmids, pKT271[170] and p171[169], featuring the ho1-pcya cassette were tested for expression of G2F and G6F chimeras from pET28c. In both cases expression of HO-1 and PcyA was induced by 1 mM IPTG. Expression was evaluated in several different *E. coli* strains, namely BL21, C41, CmpX13 and LMG194 and at different growth temperatures. While in all setups expression of G6F was too low for further studies, the combination of CmpX13 strain with the p171 chromophore plasmid grown over night at 18°C produced reasonable amounts of G2F holoprotein. After purification it showed two absorption peaks, at 530 nm and 650 nm, which were reported for the isolated *Np*R5113_GAF2 domain[165]. The equilibrium could be reversibly shifted to either side by illumination with red or green light (650 nm or 525 nm LED), which was accompanied by a visible color shift of the protein solution (Figure 12c). All further experiments involving CBCRs were hence conducted in CmpX13 and the p171 plasmid for PCB supply.

For functional testing, the CBCR-FixL fusions G2F and G6F were cloned into the pDusk plasmid[159]. In this setup the fusion protein is constitutively expressed from the $lacI^q$ promoter[171]. Histidine-kinase activity phosphorylates the encoded response regulator FixJ, which in turn drives expression of the red-fluorescent DsRed[163] from the FixK2 promoter (Figure 12d). Cultures of CmpX13 cells transformed with G2F- or G6F-pDusk and p171 were grown over night at 37°C under red or green light (25 μW cm^{-2}). Afterwards, DsRed fluorescence was measured to evaluate kinase activity under either light condition. For both constructs no kinase activity was detectable. The pDusk assay provides indirect information on the net kinase activity of the proteins (see sections 4.3 and 6.1.2 for more detailed information and discussion of bacterial assays). Still, calibrating the assay with the closely related YF1 protein and verifying expression of the fusion proteins via western blot (data not shown) suggested the low catalytic activity (low kinase or strong phosphatase activity) to be the reason for the absent DsRed signal. Lowering the

Table 7: Direct CBCR-FixL fusions.

Construct	Fusion	Activity
G6F	F2164 1037-1200 + FixL 262-505	-
G6F+2	F2164 1037-1213 + FixL 273-505	-
G2F	R5113 206-365 + FixL 260-505	-
G2F+1	R5113 206-365 + FixL 259-505	-
G2F+2	R5113 206-365 + FixL 258-505	-
G2F+3	R5113 206-366 + FixL 258-505	-
G2F-1	R5113 206-365 + FixL 260-505Δ277	-
G2F-2	R5113 206-365 + FixL 260-505Δ276-277	-
G2F-3	R5113 206-365 + FixL 260-505Δ275-277	-
G2F-4	R5113 206-365 + FixL 260-505Δ274-277	-

incubation temperature to 30°C or 18°C, like in the expression tests, did not have any observable effect.

Earlier studies on the fusion kinase YF1 showed that the length of the linker between photosensor and effector module had a great impact on the protein function[78]. Therefore, fusion variants of G2F and G6F with different linker lengths were generated. For G6F a variant, G6F+2, was created with a 40-residue linker that corresponds to the second most common pattern among GAF-HisK proteins, 7n+5 (Table 7). In G2F, single residues were systematically added or subtracted from the linker (Table 7). The deletion variants resemble modifications already successfully implemented in YF1[78]. Nevertheless, none of these adjustments led to any detectable kinase activity.

5.1.2 CcaS-FixL fusions

The consequence of the failed direct CBCR-kinase fusions was to keep the design as close as possible to natural protein architectures and the well-working YF1. The *Bj*FixL kinase is originally linked to N-terminal PAS domains, which in YF1 were replaced by a LOV domain, a subclass of PAS domains. Therefore, a CBCR featuring a GAF domain followed by a PAS domain would allow the simple exchange of one PAS domain against another. A matching candidate was the cyanobacteriochrome CcaS[172] from *Synechocystis sp.* that harbors a PCB-binding GAF domain followed by two photo-inert PAS domains and a histidine kinase (Figure 13a). The photosensor undergoes a green-red photocycle with a green-light absorbing resting state and a red-light absorbing lit state[172]. The respective fusion, denoted CF, comprises the GAF-PAS-PAS module of CcaS (residues 1-503) and the HisK of FixL (residues 258-505) joined at the prominent aspartate-isoleucine-threonine (DIT) motif that marks the end of PAS domains[78] (Figure 13a, Table 8).

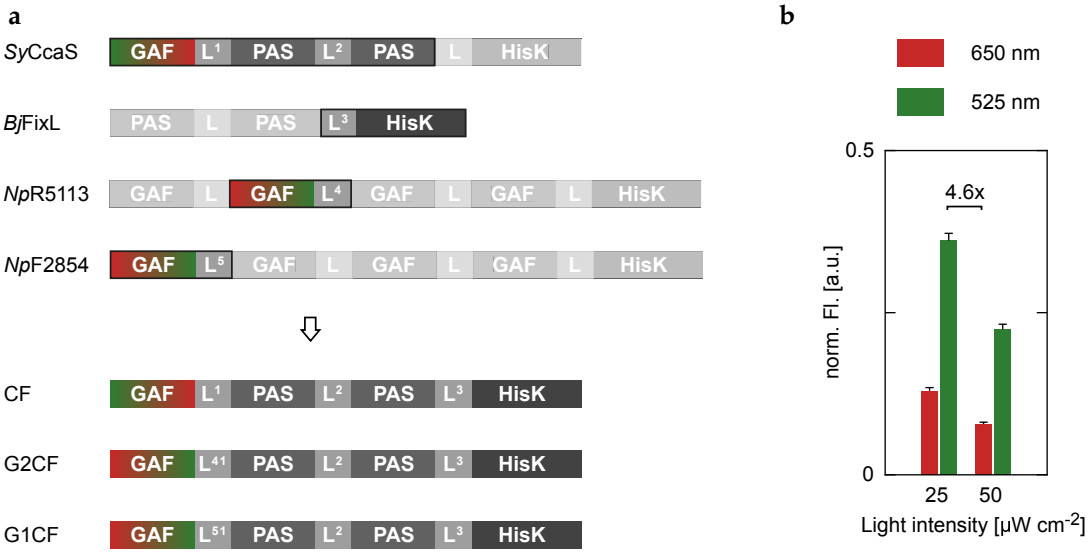

Figure 13: CcaS-FixL fusions. (**a**) Recombination of the *Sy*CcaS sensor module comprising a green-red sensitive GAF and two PAS domains with the *Bj*FixL histidine kinase yielded CF. Exchanging the *Sy*CcaS GAF for *Np*R5113_GAF2 or *Np*F2854_GAF1 led to the chimeras G2CF and G1CF. (**b**) Histidine-kinase activity of CF in the pDusk assay under red (red bars) and green light (green bars).

Table 8: Kinase activities of CcaS-FixL fusions from pDusk assay. Data are given for respective best working conditions (37°C for CF and 30°C for G2-CF and G1-CF). Light was applied at intensities of 25 μW cm^{-2} or 50 μW cm^{-2} (*).

Construct	Fusion	Activity 650 nm	Activity 525 nm	Ratio
CF	CcaS 1-503 + FixL 258-505	0.129 ± 0.005 0.078 ± 0.003*	0.36 ± 0.01 0.225 ± 0.008*	2.8 ± 0.2 2.9 ± 0.2*
CF+1	CcaS 1-503 + FixL 258-505_268vH	-	-	-
CF+2	CcaS 1-503 + FixL 258-505_268vHS	-	-	-
CF+3	CcaS 1-503 + FixL 258-505_268vHSD	-	-	-
CF+4	CcaS 1-503 + FixL 258-505_268vHSDG	-	-	-
CF-1	CcaS 1-503 + FixL 258-505Δ277	-	-	-
CF-2	CcaS 1-503 + FixL 258-505Δ276-277	-	-	-
CF-3	CcaS 1-503 + FixL 258-505Δ275-277	-	-	-
CF-4	CcaS 1-503 + FixL 258-505Δ274-277	-	-	-
CFΔPAS	CcaS 1-375 + FixL 258-505	-	-	-
CFΔ2PAS	CcaS 1-221 + FixL 270-505	-	-	-
G2CF	R5113 206-365 + CcaS 212-503 + FixL 258-505	0.71 ± 0.06	0.73 ± 0.09	1.0 ± 0.2
G1CF	F2854 1-190 + CcaS 212-503 + FixL 258-505	0.24 ± 0.02	0.22 ± 0.02	0.9 ± 0.2

In pDusk, CF showed significant and light-dependent kinase activity (Figure 13b, Table 8). After overnight incubation at 37°C DsRed levels were 2.8-fold higher under 25 μW cm^{-2} green light than under red light of the same intensity (Figure 13b). Red light of 50 μW cm^{-2} could further reduce the DsRed signal to yield a factor of 4.6. Likewise higher green-light intensities also reduced the activity to some extent. As cell viability (monitored by the optical density at 600 nm) was not affected by higher light doses, the effect probably stems from the significant green-light absorption of the red-light absorbing state of CcaS[172] causing a mixed population of both states.

To enhance the signal response of CF, I introduced linker variations described for YF1 before[78]. All of them led to completely non-functional constructs under all tested conditions (37, 30 and 18°C). Trying to reduce the size of the construct by deleting one (CF_ΔPAS) or both PAS domains (CF_Δ2PAS) yielded similar results (Table 8).

The functional CF construct gave rise to the notion that the two PAS domains could be applied as an adaptor module that mediates between GAF sensors and the kinase effector. To test this concept, in the G2CF construct the red-green *Np*R5113_GAF2 introduced above replaced the original GAF domain from *Sy*CcaS. Indeed, this fusion protein exhibited kinase activity in pDusk, but lacked light dependence (Table 8). However, in contrast to the CcaS-GAF the GAF2 from *Np*R5113 is not situated at the N-terminus of its parent protein. As the protein context can have a great influence on the function of a domain[173], I chose to switch to the red-green GAF1 of *Np*F2854

that also originally resides at the N-terminus of a protein. Nevertheless, the resulting fusion G1CF again showed kinase activity, but no light dependence.

The light-regulated CBCR-HisK fusion CF supported the notion of fusing proteins at PAS domains and suggested the *Sy*CcaS sensor as an adequate building brick for further CBCR chimeras.

5.2 Rational design of photosensitive cyclases

Based on the CcaS-FixL fusions from the previous section, the idea was to apply the same concept to engineer other homodimeric receptor proteins. Prokaryotic Type-III cyclases convert ATP or GTP into cAMP or cGMP, ubiquitous second messengers in many organisms. Like sensor histidine kinases they usually function as homodimers[174] and occur with a wide range of sensor, including the CBCR-related GAF domains.

5.2.1 Establishing a cAMP-reporter assay

There are several techniques available to assess the catalytic activity of adenylate/guanylate cyclases. The most accurate one is certainly the direct quantification of cAMP or cGMP produced in an *in-vitro* reaction via high-performance liquid chromatography (HPLC). The major drawback of this approach is that it requires purified protein, which makes testing numerous constructs rather laborious. The Miller assay represents a procedure widely used for measuring cyclase activity in cells. It relies on cAMP-dependent expression of the enzyme β-galactosidase, which in turn cleaves a reagent into a colored product[175]. Even though the assay is much faster and simpler than generating *in-vitro* data, it still has a lengthy protocol that largely hampers its applicability for testing numerous constructs and assay conditions with reasonable throughput. To address these issues I set up a bacterial fluorescence assay that does not need any additional sample treatment.

The cyclase reporter, termed CyclR, consists of the gene encoding the red-fluorescent DsRed protein under control of the lac promoter from the *lac* operon (Figure 14b). Binding of a complex of cAMP and the catabolite activator protein (CAP) induces transcription from the promoter[176]. CAP is a ubiquitous transcriptional regulator in prokaryotes and endogenous in *E. coli*. The well-characterized, light-regulated adenylate cyclase bPAC[90] served as reference to establish the assay and optimize the test conditions. A first set of experiments examined the performance of different plasmid constellations, where reporter and cyclase were either located on the same or on separate plasmids. In the two-plasmid setup bPAC was expressed from pBAD through addition of arabinose, while the reporter was situated on a pET28c (pCyclR) or pACYC184 (pCyclR_2) backbone. To simplify the system, in pBAD_CyclR both parts were assembled on one pBAD plasmid. The same argument applies to pLacI_CyclR, which is based on pDusk and puts bPAC under control of the constitutive $lacI^q$ promoter. All tests were carried out in the commercial, cyclase-deficient BTH101 strain[177] to eliminate the signal created by the intrinsic CyaA cyclase. The DsRed levels after overnight growth in liquid culture in the dark or under blue light

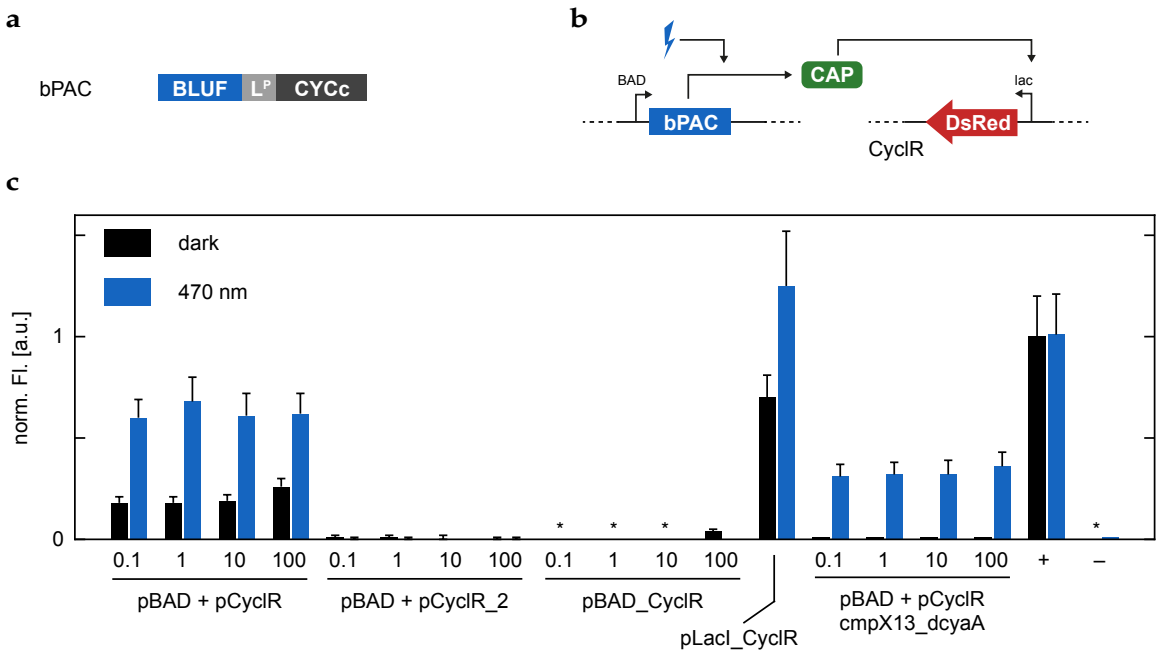

Figure 14: The CyclR adenylate-cyclase reporter. (**a**) Domain architecture of bPAC with a blue-light sensitive BLUF domain coupled to an adenylate/guanylate cyclase. (**b**) Cartoon of the CyclR reporter showing target cyclase gene under control of the BAD promoter. Produced cAMP binds the endogenous CAP protein, which induces expression of DsRed from the lac promoter. Cyclase and reporter either reside on a single or on separate plasmids. (**c**) Adenylate-cyclase activity of bPAC in the dark (black bars) and under blue light (blue bars) in different CyclR setups: bPAC was situated on pBAD and CylR on pET28c (pBAD + pCyclR) or pACYC177 (pBAD + pCyclR_2). bPAC and CyclR are united on pBAD (pBAD_CyclR) or pET28c (pLacI_CyclR). All experiments were performed in BTH101 strain. Numbers on the x-axis represent different arabinose concentrations (0.1-100 μM) for induction of bPAC expression. Best dynamic range of bPAC was observed with pBAD + pCyclR in cmpX13_dcyaA strain. pCyclR in cmpX13 served as positive control (+) and pCyclR in BTH101 as negative control (-).

(40 μW cm^{-2}) gave information on light-dependent cyclase activity. While the assay does not provide absolute catalytic activities of a construct, it allows comparing different constructs or test conditions. Here, only the combination of bPAC-pBAD and pCyclR was able to reproduce the light-induced cyclase activity reported for bPAC before[90] (Figure 14c). The ratio of light- over dark activity was roughly 3.5 and considerably lower than factors determined *in vitro*[90]. As the positive control (pCyclR in cyclase-positive CmpX13 strain) and negative control (pCyclR and empty pBAD in BTH101) indicated, especially lowering the dark signal in the assay would improve the induction ratio. Reducing the concentration of the inducing agent, arabinose, to lower the expression level of the cyclase did not show any significant improvement. Hence, in a next step I repeated the experiment in different cyclase-deficient *E. coli* strains and found the CyaA-KO of CmpX13 (provided by Roman Schubert, HU Berlin) to perform best (Figure 14c). With a low dark signal and an induction factor of 36 ± 7 the assay quantitatively represented the light-dependent cAMP production by bPAC and presented to be well suited for functional testing of cyclase proteins.

5.2.2 Cyanobacteriochrome cyclases

The CyaB1 protein from *Anabaena sp.* appeared to be a suitable target for building a chimeric CBCR cyclase[178]. It features two GAF domains, followed by a PAS domain and an adenylate/guanylate cyclase module. One of the GAF domains binds cyclic nucleotides and induces adenylate-cyclase activity in response[178]. Earlier studies already reported the engineering of *An*CyaB1 chimeras[178,179]. Although, the fusion proteins lagged behind the wild type regarding their catalytic activity, the GAF sensors were successfully swapped to change the input signal of the protein[178,179].

Here, the GAF-PAS-PAS module from *Sy*CcaS replaced the GAF-GAF-PAS module from *An*CyaB1. Like in YF1 and CF, the parent proteins were fused at their DIT motifs of the PAS domains to yield the chimera CCyc (*Sy*CcaS 1-503 + *An*CyaB1 564-859, Figure 15a). According to the preliminary tests with bPAC, CCyc was cloned into pBAD and co-transformed with pCyclR into CmpX13ΔcyaA cells. To reduce the number of plasmids in the assay, the genes for PCB synthesis were cloned onto the same pBAD backbone and put under control of a constitutive promoter[180]. However, tests in *E. coli* quickly showed that the CCyc has adenylate cyclase activity, but cannot be regulated by light (Figure 15b). At the same time Ryu *et al.* reported a similar approach and rewired a bacteriophytochrome photoreceptor to CyaB1 from *Nostoc sp.* to yield a light-dependent cyclase. The protein, termed IlaC, is activated 6-fold by red light[157]. Due to absent light regulation of CCyc and the overall low predictability of the rational design approach, further studies focused on developing alternative, more efficient engineering strategies.

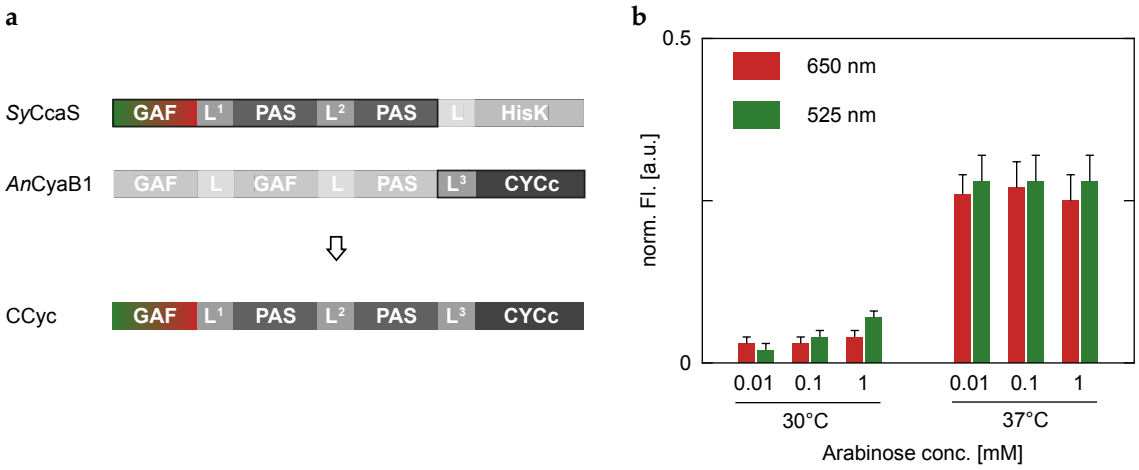

Figure 15: CBCR-coupled adenylate cyclase. (**a**) Rewiring the *Sy*CcaS sensor module to the *An*CyaB1 adenylate/guanylate cyclase yields the CCyc chimera. (**b**) Adenylate-cyclase activity of CCyc in the CyclR assay under red (red bars) and green light (green bars) upon overnight growth at 30°C or 37°C. Cyclase, CyclR reporter and PCB-synthesis genes are assembled on a pBAD plasmid.

5.3 Library-based creation of chimeric photoreceptors

The linker helix is critical for the signal propagation between sensor and effector modules and the overall protein function[71,78,156,157,181]. The above approaches to functionally rewire CBCR sensors to new effectors used rational, structure- and sequence guided design of the linker, but often failed to yield functional constructs. In principle two linkers of *n* and *m* residues yield *(n+1)·(m+1)* possible combinations. Still, only few linkers were tested for each fusion so that it remains unclear whether the impaired function stems from general incompatibility of the respective modules or the linker design. With typical linker helices consisting of roughly 20 to 100 residues it is particularly laborious to generate all fusion variants with conventional cloning. Therefore, I set up a novel strategy, named PATCHY (primer-aided truncation for the creation of hybrid enzymes), that creates fusion libraries of two protein modules with all desired combinations of their parent linkers. The protocol is inexpensive, easy to adapt and greatly accelerates the creation of numerous fusion variants.

5.3.1 Generation of YtvA-FixL libraries

Fusing the blue-light sensitive LOV domain of YtvA from *Bacillus subtilis* to the histidine kinase of FixL from *Bradyrhizobium japonicum* previously yielded the blue-light repressed kinase YF1[78] (3.3.1). Here, this fusion helped establishing and testing the PATCHY protocol. Furthermore, knowing other funtional, uncharacterized chimeras of the two modules could provide more information on the underlying signal mechanism.
The start construct, termed YF, comprised the *Bs*YtvA-LOV domain and *Bj*FixL kinase including both their full-length linkers in tandem (*Bs*YtvA residues 1-147 and *Bj*FixL residues 255-505, Figure 16a) and was cloned into the pDusk[159] reporter plasmid (Figure 16a). YF-pDusk then functioned as PCR template to create PATCHY libraries (4.4.2) using forward and reverse sets of staggered primers that anneal in the respective linker of the sensor (L^{Y}) or effector (L^{F2}). Each set contains primers with annealing sites shifted by one triplet against each other, so that they cover the entire linker sequence of the respective module (Figure 16b). The randomized linker parts were defined as residues 125-147 in *Bs*YtvA and 255-281 in *Bj*FixL and the corresponding primers were designed using a custom Python script (in collaboration with Dr. Florian Richter, HU Berlin). In short, with respect to target annealing temperature the script generates one primer for each triplet in target parts of the sensor (23 residues + 1 annealing site before the linker) and effector (27 residues + 1 annealing site after the linker) linkers. Here, 52 primers produce a theoretical library of 672 possible YtvA-FixL fusions. PCR with both sets and the YF-pDusk template generated linear plasmids and subsequent phosphorylation and ligation yielded a library containing pDusk plasmids harboring YF fusions with all chosen combinations of the parent linkers (Figure 16b).

To benchmark the PATCHY protocol, unselected libraries were analyzed by next-generation sequencing (NGS, in collaboration with Charlotte Schumacher and Prof. Dr. Andreas Möglich, HU Berlin). We thus obtained around 5,500 sequences for in-frame fusions between *Bs*YtvA and *Bj*FixL, corresponding to approximately 8-fold coverage of the theoretical library (Figure 16c).

Out of the expected 672 variants, the NGS data included sequences for 564, corresponding to a fraction of approximately 83.9 %. Data indicated that each of the primers have been used albeit to different extent. As no systematic bias toward specific lengths of fusion constructs is apparent, differences in primer usage probably stem from varying annealing/elongation efficiencies during the PCR reaction.

I generated similar libraries using the comparable ITCHY (incremental truncation for the creation of hybrid proteins)[164,182] method, which relies on incremental truncation of the starting construct with exonuclease III (also see Figure 20). A streamlined ITCHY protocol (4.4.1) served in a side-by-side comparison with PATCHY, which is discussed in detail in section (6.2.1).

The libraries were screened for functional constructs using the pDusk reporter. Cultures of *Escherichia coli* DH10b were transformed with ITCHY or PATCHY libraries and grown on agar plates under blue light or in the dark. Colonies harboring functional kinase fusions were identified based on their DsRed fluorescence and isolated for further investigation. After growing

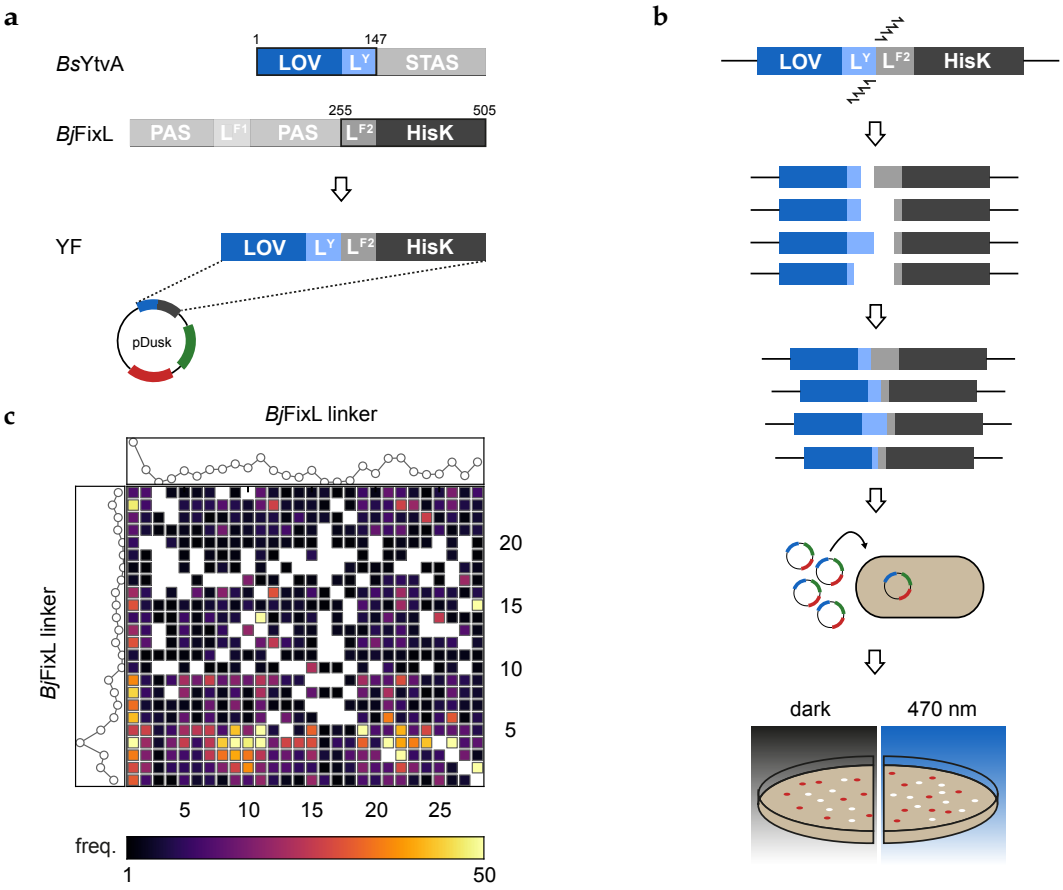

Figure 16: YtvA-FixL PATCHY libraries. (**a**) The starting construct YF comprises a tandem of the *Bs*YtvA-LOV domain and the *Bj*FixL histidine kinase and both full-length linkers cloned into pDusk. (**b**) PCR with staggered primers that anneal in the linker regions yield linear truncated plasmids. Subsequent phosphorylation and ligation generates PATCHY libraries containing random combinations of the parental linkers. Libraries are transformed into *E. coli* for functional testing using the pDusk assay. (**c**) Library coverage determined by next-generation sequencing. Numbers on the x-axis represent the 28 primers annealing in the effector linker. Numbers on the y-axis represent the corresponding 24 primers for the sensor linker. Line graphs indicate usage of each primer and color-coding denotes the total copy number of each construct in the NGS sample.

the selected clones in 96-well liquid culture in the dark and under blue light, the resulting DsRed fluorescence represented the light-dependent activity of the respective constructs.

5.3.2 Function of YtvA-FixL fusion variants

This section combines the results from both, ITCHY and PATCHY, libraries. Table 9 lists all characterized clones and marks their origins. Apart from a large fraction of constitutively active variants the libraries contained a multiplicity of light-switchable YtvA-FixL fusions with various linker lengths. From the theoretical 672 variants, 29 were identified as light-regulated histidine-kinases in pDusk (Figure 17a). Here, the linker length was counted from residue 125 of *Bs*YtvA to the conserved histidine of *Bj*FixL (residue 291). Consequently, YF1 that comprises the residues 1-127 from *Bs*YtvA and 258-505 from *Bj*FixL, has a linker length of 37. As anticipated the libraries contained YF1 as well as the previously reported YF2 (*Bs*YtvA 1-135 + *Bj*FixL 266-505) and YF4 (*Bs*YtvA 1-147 + *Bj*FixL 278-505)[78], which feature the same linker length, but a different linker sequence compared to YF1. As expected, these constructs showed reduction of kinase activity by blue light[78,159] (Figure 17b, Table 9). The rather small fraction (< 5%) of light-dependent variants displays how much signal propagation across the linker depends on its amino-acid composition. Changes of single residues in the linker can abolish signal propagation[67,160] or even invert its polarity[78,157,181]. While YF_L58 shows strong reduction of kinase activity upon blue-light illumination, in YF_L59 an additional alanine in the linker leads to a signal inversion and 2-fold induction of kinase activity by light. However, with the same linker length, but by changing one single residue +21_T148A and +22_A148L completely lack light sensitivity (Figure 17a, Table 9).

Looking at the collective sequences of the light-switchable constructs in this study it occurs that all except one (YF_L15) retain the residues D125, I126 and T127 of *Bs*YtvA (Figure 17a). This DIT motif is highly conserved among LOV and other PAS domains[63,78,166] and apparently crucial for proper signal transduction in the protein. With the shortest linker of a light-switchable construct being 13 residues and the longest 59 residues, the total distance between sensor and effector module seems to be of minor importance for signal propagation along the helix. Nevertheless, light regulation primarily occurs at specific linker lengths. While YF1, YF2 and YF4 feature linkers of 37 residues; other light-repressed fusions have linkers that are 7, 14 and 21 residues shorter or longer, ergo have linker lengths of 7n+2 (Figure 17a, b). Structural data on YF1 suggest that the helical linkers form a coiled coil in the functional homodimer[67]. In a coiled coil the helix angle changes by roughly 103° per residue, meaning that seven residues correspond to two complete helical turns and conserve the relative orientation of the attached sensor and effector modules, i.e. the LOV domain and histidine kinase. Plotting the helix angle against the respective number of identified light-regulated YtvA-FixL fusions shows that one angle, here defined as 0°, is clearly favored among light-repressed fusions (Figure 17c). The same is true for light-induced constructs with the difference that the favored angle is shifted by 103° so that these constructs typically have linkers of 7n+3 residues. Nevertheless, exceptions from these rules occur for both, light-induced and repressed variants. Variants like YF_L53 (7n+4) however completely break the rule and point at the importance of the linker sequence. In the case of YF_L53 a spontaneous mutation created a

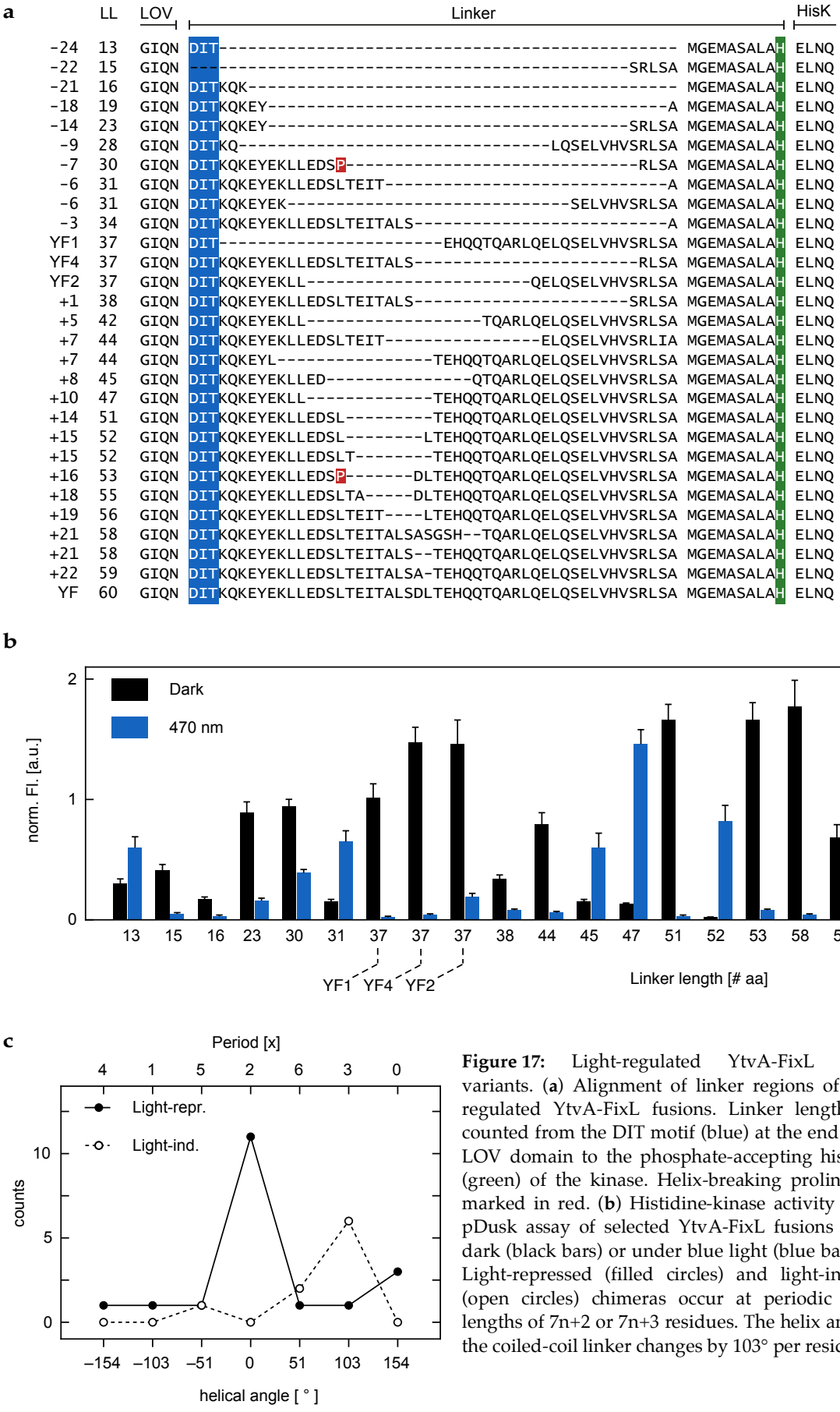

Figure 17: Light-regulated YtvA-FixL fusion variants. (**a**) Alignment of linker regions of light-regulated YtvA-FixL fusions. Linker lengths are counted from the DIT motif (blue) at the end of the LOV domain to the phosphate-accepting histidine (green) of the kinase. Helix-breaking prolines are marked in red. (**b**) Histidine-kinase activity in the pDusk assay of selected YtvA-FixL fusions in the dark (black bars) or under blue light (blue bars). (**c**) Light-repressed (filled circles) and light-induced (open circles) chimeras occur at periodic linker lengths of 7n+2 or 7n+3 residues. The helix angle in the coiled-coil linker changes by 103° per residue.

Table 9: Light-repressed (top) and light-induced (bottom) YtvA-FixL fusions from ITCHY and PATCHY libraries. * marks variants exclusively isolated from ITCHY libraries.

Construct	Linker length	Periodicity	Activity dark	Activity 470 nm	Ratio
YF_L15	15	1	0.41 ± 0.05	0.05 ± 0.01	8 ± 3
YF_L16	16	2	0.17 ± 0.02	0.03 ± 0.01	5 ± 2
YF_L19	19	5	1.5 ± 0.1	0.54 ± 0.06	2.8 ± 0.5
YF_L23	23	2	0.89 ± 0.09	0.16 ± 0.02	6 ± 1
YF_L28	28	0	0.76 ± 0.07	0.13 ± 0.02	6 ± 1
YF_L30*	30	2	0.94 ± 0.06	0.39 ± 0.03	2.4 ± 0.3
YF_L34	34	6	0.18 ± 0.02	0.057 ± 0.007	3.2 ± 0.7
YF1	37	2	1.0 ± 0.1	0.021 ± 0.006	50 ± 20
YF4	37	2	1.5 ± 0.1	0.043 ± 0.005	34 ± 7
YF2	37	2	1.5 ± 0.1	0.19 ± 0.03	8 ± 2
YF_L38*	38	3	0.34 ± 0.03	0.08 ± 0.01	4.3 ± 0.9
YF_L42	42	0	0.26 ± 0.03	0.032 ± 0.008	8 ± 3
YF_L44-1	44	2	0.8 ± 0.1	0.06 ± 0.01	13 ± 4
YF_L44-2	44	2	0.89 ± 0.06	0.22 ± 0.02	4.1 ± 0.6
YF_L51	51	2	1.7 ± 0.1	0.034 ± 0.005	50 ± 10
YF_L53*	53	4	1.7 ± 0.1	0.084 ± 0.009	20 ± 4
YF_L56	56	0	0.27 ± 0.04	0.018 ± 0.004	15 ± 5
YF_L58-1	58	2	1.8 ± 0.2	0.040 ± 0.007	40 ± 10
YF_L58-2	58	2	1.3 ± 0.2	0.018 ± 0.005	80 ± 30
YF_L58 T148A*	58	2	1.8 ± 0.2	2.1 ± 0.2	0.9 ± 0.2
YF_L13	13	6	0.30 ± 0.04	0.60 ± 0.09	0.5 ± 0.1
YF_L31-1	31	3	0.014 ± 0.003	0.10 ± 0.02	0.14 ± 0.06
YF_L31-2	31	3	0.15 ± 0.02	0.65 ± 0.09	0.23 ± 0.07
YF_L45	45	3	0.15 ± 0.02	0.6 ± 0.1	0.24 ± 0.08
YF_L47	47	5	0.13 ± 0.01	1.5 ± 0.1	0.09 ± 0.02
YF_L52-1	52	3	0.016 ± 0.004	0.8 ± 0.1	0.02 ± 0.01
YF_L52-2	52	3	0.005 ± 0.003	0.21 ± 0.03	0.03 ± 0.02
YF_L55	55	6	0.067 ± 0.007	0.71 ± 0.07	0.09 ± 0.02
YF_L59*	59	3	0.7 ± 0.1	1.6 ± 0.3	0.4 ± 0.1
YF_L59 A148L*	59	3	1.8 ± 0.2	2.0 ± 0.2	0.9 ± 0.2

proline, which disturbs the helix integrity and the periodic order of hydrophobic residues typical for coiled coils (Figure 17a, b, Table 9).

5.3.3 Generation and screening of YtvA-cyclase libraries

The PATCHY strategy provided a tool to revisit the engineering of photoactive cyclases. In a first step, the *Bs*YtvA-LOV domain applied in the previous section, was supposed to replace the GAF-GAF-PAS module in *An*CyaB1. As LOV domains are a subtype of PAS domains, this appeared to be a straightforward approach. The corresponding start construct, YCy (*Bs*YtvA 1-147 + *An*CyaB1 564-859, Figure 18a), was cloned into pBAD and PATCHY libraries were created as described before (5.3.1). However, in the absence of a cyclase-deficient cloning strain, all cyclase libraries were first transformed into *E. coli* DH10b to yield a sufficient number of transformants. After overnight growth on agar plates approximately 10.000 colonies were pooled and plasmid DNA was isolated. This library was then co-transformed with the pCyclR reporter into CmpX13_ΔcyaA for functional testing. Library screening did not reveal any light-switchable YtvA-CyaB1 fusion. It turned out to be a major drawback of *An*CyaB1 that the isolated cyclase module retained significant activity. Consequently, most fusion constructs showed robust cyclase activity hampering discrimination of light-switchable constructs from the background.

I switched to the bPAC cyclase to circumvent this issue. It comprises a blue-light sensitive BLUF domain and an adenylate cyclase, which are connected by a linker, which presumably forms a helical structure. The conserved overall topology of BLUF and LOV domains suggested that they could also be functionally interchangeable. Thus, I applied the PATCHY protocol to the YP fusion of the YtvA-LOV domain (residues 1-147) and the bPAC cyclase (residues 95-350) (Figure 18a). Indeed, most fusions appeared to be inactive in the functional assay. From the number of active clones four were finally identified as light-induced (Figure 18b, Table 10). The induction factors of two to threefold are low, but could be reproduced *in vitro* (unpublished data, Charlotte H. Schumacher, HU Berlin).

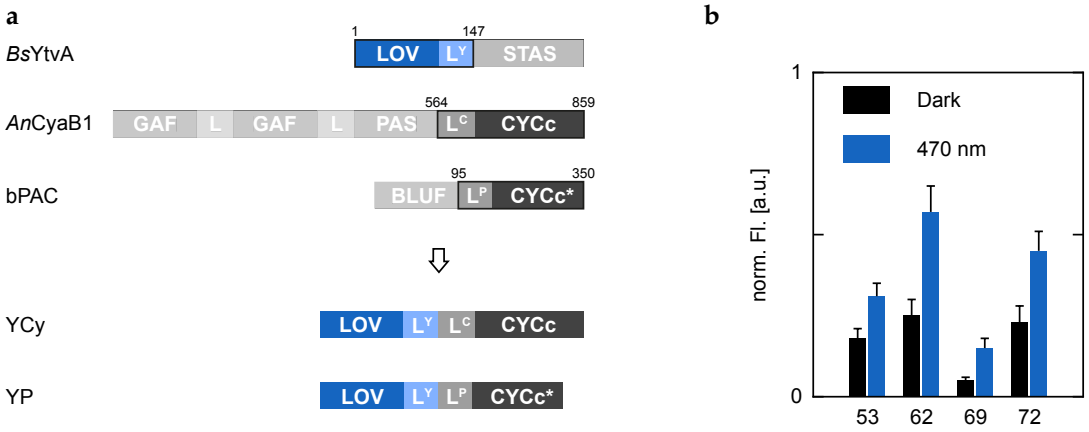

Figure 18: YtvA-cyclase PATCHY libraries. (**a**) Starting constructs for fusion libraries comprise a tandem of *Bs*YtvA-LOV and either the adenylate cyclase module from *An*CyaB1 or bPAC cloned into pBAD. (**b**) Adenylate-cyclase activity of YtvA-bPAC chimeras from the pCyclR assay in the dark (black bars) and under blue light (blue bars).

Table 10: Light-switchable YtvA-bPac fusion variants from PATCHY libraries.

Construct	Linker length	Periodicity	Activity dark	Activity 470 nm	Ratio
YP-L53	53	4	0.18 ± 0.03	0.31 ± 0.04	0.6 ± 0.2
YP-L62	62	6	0.25 ± 0.05	0.57 ± 0.08	0.5 ± 0.1
YP-L69	69	6	0.05 ± 0.01	0.15 ± 0.03	0.3 ± 0.1
YP-L72	72	2	0.23 ± 0.05	0.45 ± 0.06	0.5 ± 0.2

5.4 Linker analysis of natural receptor protein

Heptad periodicity of linker lengths in YtvA-FixL fusions[78] in the previous section (5.3.2) suggested a critical role of the relative angular orientation of sensor and effector modules. Alignments of natural PAS-histidine kinases revealed that the respective linker lengths generally also group into classes differing by seven residues[78]. Here, I asked whether other sensor-effector combinations comply with a similar pattern.

I identified and aligned the amino-acid sequences of natural proteins comprising PAS, GAF, BLUF or phytochromes (PHY) sensor modules that are directly coupled to C-terminal histidine kinase (HisK), adenylate/guanylate cyclase (CYCc), diguanylate cyclase (GGDEF) or phosphodiesterase (PDE) effector modules and evaluated linker lengths among them. While BLUF and PHY exclusively occur in photoreceptors, PAS and GAF also include sensors of redox potential or small metabolites. The chosen effectors represent common module types among homodimeric signal receptors that include photosensors so that they are particularly attractive for engineering optogenetic tools.

Table 12 summarizes the respective data. A closer look on the linker sequences reveals that linkers of almost all identified cyclase proteins (BLUF-CYCc, PAS-CYCc, GAF-CYCc) contain multiple prolines. The same is true for PAS-PDEs. In contrast, linkers of HisK and GGDEF proteins are virtually proline-free (Table 11). Some sensor-effector combinations, i.e. BLUF-GGDEF, BLUF-HisK, BLUF-PDE, PHY-CYCc and PHY-PDE, did not match any protein in the Pfam database and were consequently neglected. PAS-PDE, GAF-PDE, BLUF-CYCc and PHY-GGDEF did only match few sequences, which prohibits statistical analysis. Furthermore, poor alignment quality impeded evaluation of heptad periodicity in PHY-HisK proteins. The remaining combinations GAF-HisK, PAS-HisK, GAF-GGDEF, PAS-GGDEF, GAF-CYCc and PAS-CYCc show significant heptad periodicity of their linker lengths (Figure 19). Consequently, conserved angular orientation of sensor and effector modules is a common theme among various signal-receptor proteins. Section 6.3.2 further discusses implications for the signaling mechanism and module interchangeability.

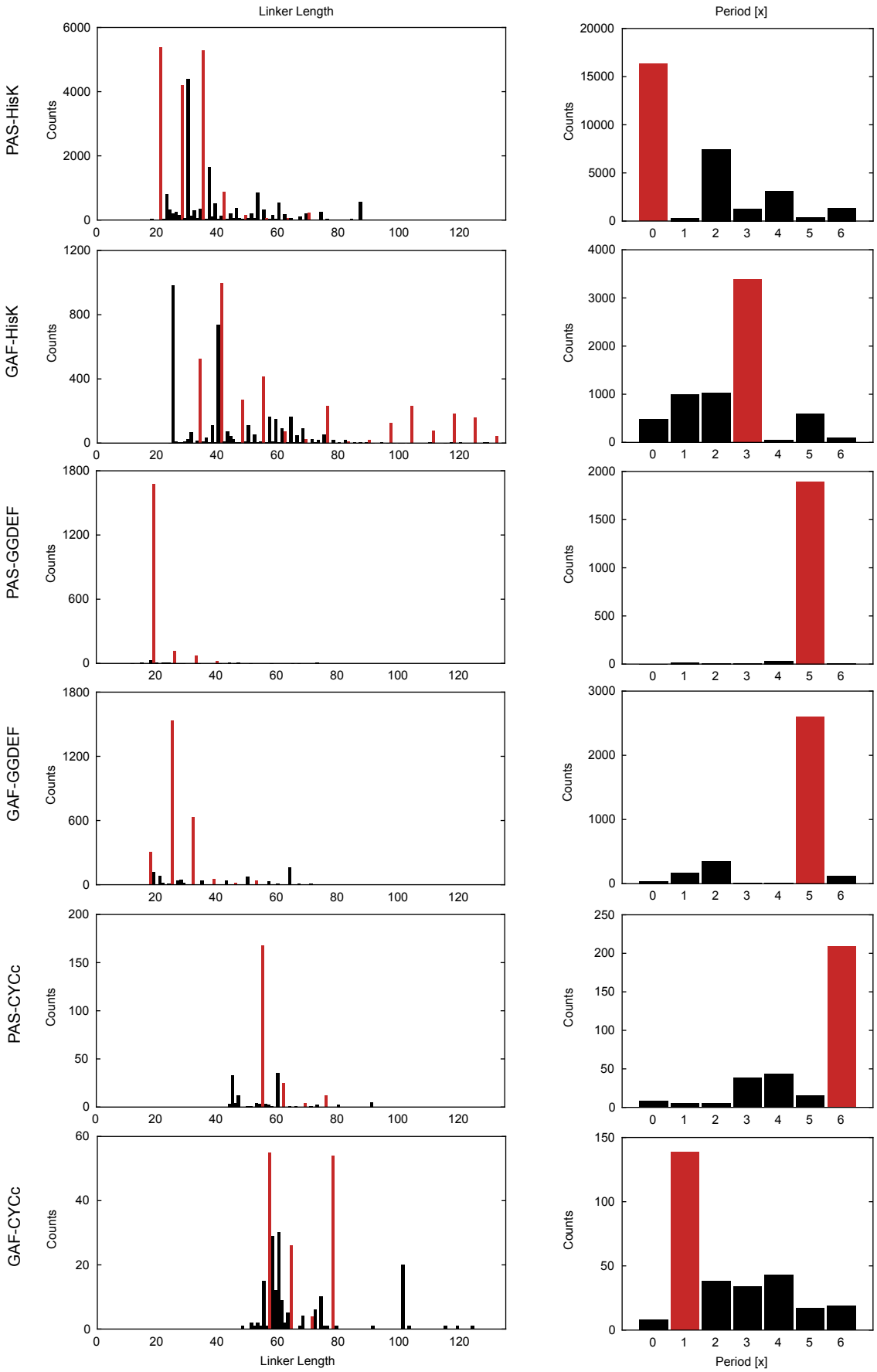

Figure 19: Linkers of natural signal-receptor proteins. The histograms show linker-length distributions (left) and heptad periodicities (7n+x, right) of PAS-HisK, GAF-HisK, PAS-GGDEF, GAF-GGDEF, PAS-CYCc and GAF-CYCc proteins. The dominant heptad period (x) and corresponding linker lengths are marked in red.

Table 12: Linker analysis of natural receptor proteins. Heptad periodicities of linker lengths occur for sensor-effector combinations marked with '+'. Pairs marked with '-' did match none or few entries in the Pfam database and were neglected for further analysis. The '?' denotes cases where poor alignment quality impeded proper analysis of the heptad periodicity.

	PAS	GAF	PHY	BLUF
HisK	+	+	?	-
CYCc	+	+	-	-
PDE	-	-	-	-
GGDEF	+	+	-	-

Table 11: Prolines in the linkers of natural signal receptors. Numbers represent the number of prolines in the linker divided by the total number of proteins from the respective sensor-effector combination.

	PAS	GAF	PHY	BLUF
HisK	0.14	0.04	0.02	-
CYCc	1.44	1.75	-	4.65
PDE	-	3.51	-	-
GGDEF		0.03	0.03	-

6 Discussion

6.1 Engineering cyanobacteriochrome chimeras

6.1.1 Rational design

Compared to most other sensory photoreceptors cyanobacteriochromes were only recently found and are hence relatively uncharted. Their photosensor modules adopt a GAF fold, which is widespread among cytosolic photo- and chemoreceptors and occurs in combination with various effectors like histidine kinases, cyclases or MCP domains[183]. CBCRs are spectrally diverse, bistable switches, so that engineering does not only contribute to mechanistic understanding of signal transduction between the CBCR sensor and respective effector modules, but could also yield novel optogenetic tools.

To my knowledge, there are no studies published on engineering chimeric CBCR proteins, yet. Therefore, the idea was to first rewire a CBCR sensor to a well-characterized effector module of a class that also occurs in natural CBCR proteins. The *Bj*FixL histidine kinase meets these requirements and was previously employed in the functionally[63] and structurally[67] well characterized YF1 chimera. The chimeras G2F and G6F employed the CBCRs *Np*R5113_GAF2 and *Np*F2164_GAF6, which exhibit good expression in *E. coli*, and were designed with respect to sequence and structure analysis (5.1.1). Still, both fusion proteins did not exhibit any histidine-kinase activity in the pDusk assay and variation of the linker between sensor and effector module, which led to light-regulated constructs in other photoreceptor-engineering approaches[157,158], failed to improve the function of either fusion.
Various reasons could account for these findings. First of all, the two GAF domains could elicit a light-dependent structural change that is not compatible to the *Bj*FixL kinase. Moreover, both sensors originate from CBCRs with multiple GAF domains, which raises the question whether the single GAFs can generally function outside their protein context. Although each GAF domain binds its own chromophore and is photochemically active[184], signaling to the effector potentially only works in conjunction with neighboring sensor modules. PAS/LOV domains feature the highly conserved DIT motif that marks the end of the domain core β sheet and the linker helix (3.1.1). In YF1, fusing *Bs*YtvA and *Bj*FixL at their DIT motifs conserved the complete FixL linker. GAF domains do not have such a clear boundary and end in an α helix that interacts with the domain core (3.1.2, Figure 3c). Hence, the fusion site has to be further down the helix and the linkers of both parent proteins have to be truncated and fused. Arguably, this process is more error-prone and easily produces irregularities in the linker helix impairing intramolecular signal transduction. Further systematic engineering of cyanobacteriochromes will identify, which of the discussed explanations is actually valid.

Using the sensor module of CcaS from *Synechocystis sp.* circumvents these problems (5.1.2). It comprises only a single GAF domain and two photo-inert PAS domains so that the complete, natural sensor module is conserved and fused at a DIT motif to keep the *Bj*FixL linker. Indeed,

the resulting chimera CF exhibits light-regulated histidine-kinase activity. The CcaS-GAF domain has a green-red photocycle and CF shows high net kinase activity in the pDusk assay under green light that is reduced by up to 4.6-fold upon red-light illumination. Structural data on CF would unravel how the signal from the GAF domain propagates through two PAS domains and which signal-induced structural changes occur in the kinase. Comparing structural data and the effects of linker variation or mutation in YtvA-FixL and CcaS-FixL fusions could reveal similarities and differences of signaling between the two sensors and the kinase effector. First experiments did not detect any kinase activity of CF linker variants in pDusk (5.1.2). *In-vitro* experiments reported that the corresponding linker variants were functional in YtvA-FixL, but caused significantly lower catalytic activities[78]. Based on the overall lower output signal of CF in pDusk, activities of the linker variants were possibly below the detection limit of the assay. Thus, in a next step systematic linker variation with PATCHY described for YtvA-FixL (5.3.1) should be applied to CcaS-FixL.

CF provided the foundation for further engineering. Deletion of one or both PAS domains was supposed to reduce the size and complexity of the construct, but yielded inactive chimeras (5.1.2). Keeping the PAS domains as adaptors to couple other CBCR sensors to the kinase produced functional, but light-insensitive proteins (5.1.2). Again, similar arguments apply as mentioned above: Either the tested CBCR sensors do not work outside their original protein context or the linker design was disadvantageous. The G2CF construct featured the *Np*R5113_GAF2 domain that originally resides between two GAF domains. Using the N-terminal *Np*F2854_GAF1 in G1CF was thought to solve this problem, still failed to yield a light-regulated protein. Yet, GAF1 misses the original three GAF domains from the parent protein, which could be critical for functionality. Taken together, the FixL fusions indicated that first, the CcaS-sensor module works well under the applied test conditions in *E. coli*. Second, they emphasize that PAS domains are convenient building bricks, because their conserved DIT motifs clearly mark the boundary between the domain-core β sheet and the linker helix. This bolsters the notion that fusing target proteins by exchanging PAS domains might be a valid strategy to engineer other photoreceptors[78].

Following up on that, CyaB1 from *Synechocystis sp.* provided a natural target for engineering CBCR-coupled adenylate cyclases. Its N-terminal sensor module comprising two GAF and one PAS domain was replaced by the CcaS-sensor module by fusing the two proteins at their DIT motifs (5.2.2). However, the chimera did not show any light dependence in the bacterial reporter assay. Whether this stems from general incompatibility of the CcaS sensor and the CyaB1 cyclase or linker variation can restore light regulation remains to be investigated. This fusion displays that designs closely resembling the original protein architecture are still prone to error. As discussed for CBCR-FixL fusions above, again little information can be drawn from a handful of non-functional chimeras. Instead, numerous module combinations, i.e. diverse CBCRs and effectors, from different protein architectures have to be tested in order to derive reliable knowledge on compatibility and shared signaling mechanisms. Moreover, various linker lengths and sequences have to be implemented to minimize the possibility of non-functional constructs due to faulty linker designs.

6.1.2 Protein activities in bacterial assays

The dynamic range of CF in pDusk agrees well with studies on the two-component system of *Sy*CcaS and the cognate response regulator CcaR[156,185,186]. Yet, bacterial assays employed in this study do not allow deriving absolute protein activities. How kinase activity quantitatively translates to the assay output depends on the respective test system. Systematic optimization of expression levels in the CcaS-CcaR system increased the respective dynamic range from roughly threefold[185] to over 100-fold[180]. Therefore, tuning the expression levels of CF, the response regulator FixJ and the PCB-providing enzymes HO-1 and PcyA in the pDusk-reporter plasmid might increase the dynamic range to yield a robust, bistable gene-expression system regulated by red and green light. However, this demonstrates that dynamic ranges are only comparable for different proteins in the same bacterial test setup. Due to the cooperative behavior of many receptors and enzymes, signaling networks often have non-linear dose-response curves. Hence, small changes in an input may elicit large changes of the output signal and *vice versa*. Not necessarily the actual dynamic range of the protein, but its fine-tuning to the signaling network determines whether it performs well enough to trigger target output signal.

The pDusk (3.3.1) as well as the pCyclR (5.2.1) reporters facilitate quick testing of chimeric histidine kinases and adenylate cyclases under various experimental conditions. In the case of pDusk, YF1 demonstrates that the assay is able to display the function of FixL-fusion proteins[159]. Furthermore, light-dependent activity of CF indicated sufficient supply of the PCB chromophore for testing CBCR-FixL chimeras (5.1.2). The pCyclR assay readily reproduced light-dependent cyclase activity of bPAC (5.2.1) and indicated cAMP production by the CcaS-CyaB1 chimera. Although, the applied PCB-synthesis cassette was previously reported to produce functional CcaS holoprotein in *E. coli*[180], it is possible that PCB supply was insufficient in the present test context. Western-blot analysis and fluorescence outputs suggested expression of the fusion proteins. Still, in pDusk and pCylR assays moderate expression levels of respective fusion proteins impede quantifying chromophore incorporation via gel electrophoresis making it difficult to assess exact amounts of functional holoprotein. In contrast to *in vitro* tests, protein expression and functional testing both happen in parallel and experimental parameters like incubation temperature and duration, light doses, inducer concentration or plasmid combinations directly affect both processes. This makes the test system rather complex and complicates distinguishing between non-functional protein designs and insufficient expression of holoprotein in the test system. Moreover, poorly designed proteins often do not fold properly and hence yield low amounts of holoprotein. While *in-vitro* approaches provide well-defined systems, where single parameters like reactant concentrations can be easily controlled, the optimization of bacterial assays is more challenging. Hence, not much information derives from non-functional fusion proteins unless larger numbers of different designs are tested. Here lies the major problem of the rational-design approach. In the present study this strategy did not reliably produce light-regulated chimera proteins. Instead the results suggest the 'hit ratio' of structure- and sequence-guided design to be rather low in this context. Therefore, on the one hand numerous fusion proteins have to be tested to yield a light-regulated chimera, which requires a fast (hence *in vivo*)

assay. On the other hand, if most of these variants show no or impaired function, these assays make the design process resemble 'fishing in troubled water'.

In conclusion, rational design proved to be inefficient in the present context. Instead, module recombination in homodimeric signal receptors requires systematically testing large numbers of fusions to create light-regulated chimeras and deduce information on general design principles from functional constructs. Indeed, bacterial assays lend themselves to high-throughput screening of various sensor-effector combinations. The following sections discuss alternative strategies for module rewiring (6.2.2) and conclude specific approaches for engineering cyanobacteriochromes (6.4.1) and other receptor proteins (6.4.2).

6.2 Library-based creation of chimeric photoreceptors

6.2.1 The PATCHY strategy

Preceding sections illustrated difficulties of rewiring protein modules in homodimeric photoreceptors. Previous designs of homodimeric photoreceptor chimeras reported testing only few linker variants to create a light-regulated protein[157,158,78,156]. But these studies might not represent the average case. As usually hundreds or thousands possible combinations of the parental linker exist, one can expect many studies to be canceled after cloning and testing a handful of linkers and the results to remain unpublished.

The PATCHY approach (5.3.1) creates fusion libraries of two protein modules containing all desired combinations of their parent linkers at once and renders cloning single variants obsolete. It uses PCR with two sets of staggered primers and a template comprising the sensor and effector modules and their full-length linkers cloned in tandem. The forward primers anneal to different positions within the C-terminal effector module and are each shifted in their annealing sequence by one triplet. The same applies to the reverse primers annealing in the linker of the N-terminal sensor module. The PCR reaction linearizes the plasmids and randomly truncates the two linkers. Phosphorylation and ligation creates the final fusion library. The protocol was applied to a fusion of the *Bs*YtvA-LOV domain and the histidine kinase from *Bj*FixL that was cloned in the pDusk reporter plasmid. In general for two linkers with n and m residues, $n+m+2$ primers generate $(n+1)\cdot(m+1)$ fusion variants. Here, target linker regions comprised 23 residues (125-147) in *Bs*YtvA and 27 residues (255-281) in *Bj*FixL so that it needs 52 primers to create the library of 672 chimeras. As all primers are applied in a single PCR reaction they were designed to have roughly similar annealing temperatures. Sample reactions with single primer pairs showed that the script generates functional primers. Also library screening revealed YtvA-FixL fusions of various linker lengths and compositions. Section 6.3.1 discusses the function of these constructs in more detail. Sequencing a random ensemble of *Escherichia coli* clones (in collaboration with Prof. Dr. Andreas Möglich at Universität Bayreuth and Charlotte Schumacher at HU Berlin) transformed with PATCHY libraries showed that all primers worked in the PCR and contributed constructs to the library (Figure 16c). While there is no systematic bias toward YtvA-FixL fusions of specific

lengths, it should be possible to modify the construct distribution by adapting primer concentrations. Based on sequencing results the PATCHY reaction can be repeated with a subset of the primers to specifically generate linker variants that were initially underrepresented. Alternatively, a non-uniform primer distribution might be chosen right from the start to deliberately bias the resultant PATCHY libraries toward certain constructs or groups of constructs.

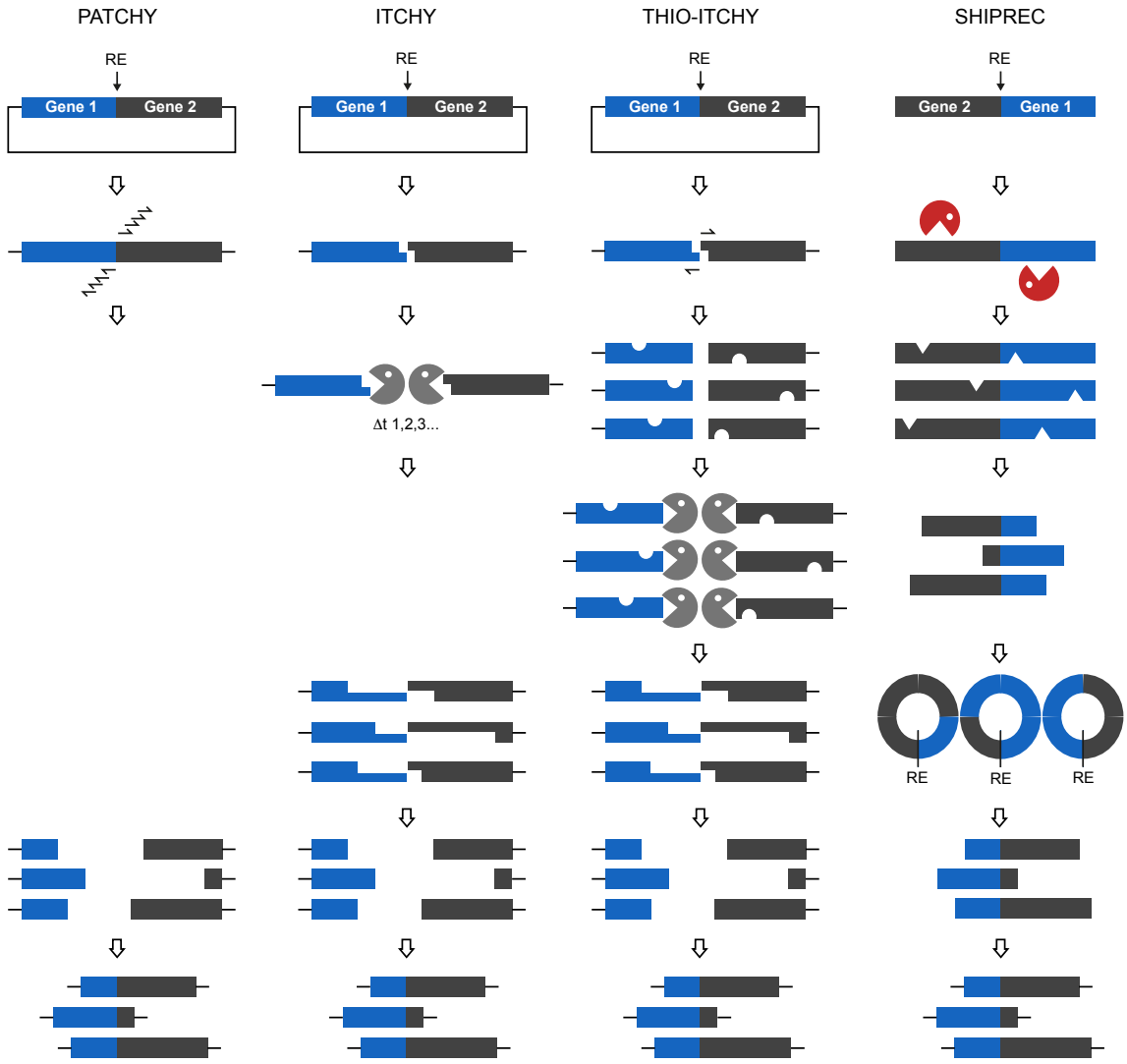

Figure 20: Methods for generating chimera libraries. ITCHY and PATCHY start out from the same template plasmid that comprises two target genes cloned in tandem with an interjacent unique restriction site (RE). **PATCHY** uses PCR with staggered forward and reverse primers to create linear plasmids with genes 1 and 2 truncated at position corresponding to the primer annealing sites. Phosphorylation and ligation yields a library of plasmid harboring target gene chimeras. **ITCHY** involves linearization of the template plasmid using the RE site. Next, exonuclease incrementally truncates 5′ ends of both genes. Time-point sampling and blunt-end generation with S1 nuclease produces randomly truncated linear constructs that are ligated to yield the fusion library. **THIO-ITCHY** uses PCR amplification of the linearized template to randomly incorporate nucleotide analogs. Exonuclease digest stops at the first incorporated analog to yield recessed 5′ ends. Subsequent steps are similar to ITCHY. **SHIPREC** starts from a fusion gene with inverse orientation to the final, target chimeras. DNaseI randomly nicks the DNA and S1 nuclease cleaves the DNA at the nick sites to yield randomly truncated fusion genes. Blunt-end ligation and linearization using the RE site generates fusion variants with the correct orientation of gene 1 and 2. Cloning the random gene fusions into a plasmid creates the final libraries.

A comparable approach for accomplishing similar module fusions is ITCHY. Like PATCHY, it generates fusion libraries in a sequence-homology independent manner. Starting out from the same template as PATCHY, it involves linearization of the plasmid at the fusion site and subsequent incremental truncation by exonuclease III. Stopping the reaction at different time points and subsequent blunt end ligation leads to a library of random gene fusions[164,182,187] (Figure 20). A variation of the protocol, termed THIO-ITCHY, uses the random incorporation of α-phosphothioate nucleotides, which stop the exonuclease digestion and render the time-point sampling obsolete[164,188]. PCR amplification of the linearized starting construct facilitates the incorporation (Figure 20). Subsequent exonuclease digest halts at the first incorporated nucleotide analog and S1 nuclease cleaves the resulting overhangs yielding the randomly truncated construct. SHIPREC applies random truncation of a gene fusion by DNase digest[189] (Figure 20). Of all approaches, the original ITCHY protocol, using time-dependent exonuclease digest, produces the most uniform distributions of fusion constructs[190]. Furthermore, it does not need unconventional nucleotide analogs. ITCHY libraries of YtvA-FixL served in a side-by-side comparison with PATCHY.

Compared to PATCHY, ITCHY, THIO-ITCHY and SHIPREC have lengthy protocols involving multiple enzyme reactions that have to be tweaked. Although, the ITCHY protocol was streamlined using enzymes that work in the same buffer or in a mix, several additional purification steps are necessary. Therefore, especially ITCHY and SHIPREC, which do not involve PCR, require large DNA amounts. As speed and processivity of the exonuclease are hard to control, ITCHY generates significant proportions of unwanted byproducts. In my hands extensive optimization of reaction time, reaction temperature, enzyme and template concentrations finally resulted in libraries containing desired linker variants (5.3.2). Yet, random sampling showed that the amount of overly truncated constructs could be reduced, but was still very high. It is hardly possible to restrict the random digestion to short DNA stretches like the linker regions. THIO-ITCHY lacks the time-dependent exonuclease digestion, but cannot fully circumvent this problem. Tuning the incorporation rate of nucleotide analogs precludes large amounts of highly truncated constructs, but fails to generate homogeneous distributions of desired constructs[190]. Moreover, all ITCHY versions fall to the fact that even in the ideal case, two thirds of the library feature frame-shifted triplet pattern. SHIPREC has an even longer and more complex protocol than ITCHY, with additional cloning steps (Figure 20). Similar to ITCHY, SHIPREC is not triplet sensitive and creates significant fractions of frame-shifted chimeras.

Herein lies a major advantage of the novel PATCHY strategy, because oligonucleotide primers allow restricting the random fusion to defined parts of the template construct, e.g. the linker regions. Additionally, they minimize frame shifts. Hence, PATCHY provides an efficient method for generating libraries of defined gene fusions with a quick and easily implemented protocol.

6.2.2 Application of PATCHY for linker scanning

Rational design of CBCR chimeras underlined the need for comprehensive and systematic recombination of protein modules to assess functional compatibility. Sensors and effectors from different protein architectures, e.g. CBCR sensors from single GAF or multi-domain

arrangements, have to be tested with various linker designs. PATCHY libraries especially tackle the linker problem by providing faster construct design. In a brute-force approach it generates all possible linker combinations and uses high-throughput screening to isolate the functional variants from the library. This linker scanning is applicable for mechanistic studies as well as building chimeric proteins, e.g. for optogenetic applications. Section 6.4 discusses possible targets for further engineering of molecular tools. Methodical testing of all linker variants maps how the linker sequence relates to the protein function and allows identifying critical residues or motifs and general connectivity rules. Hence, library-based approaches are complementary to rational design in the sense that instead of applying mechanistic knowledge to engineer chimeras, they allow deducing information from full sets of fusion variants.

Testing large numbers of possible linker combinations can also guide the design process of molecular tools, because it accelerates the identification of functional constructs and the evaluation of whether two particular modules are functionally compatible at all. Furthermore, previous designs usually derived from testing few variants so that better working fusion variants might exist. Indeed, engineered chimeras often lack behind the parent protein regarding catalytic activity or dynamic range[178,179,157]. In fact, the same argument might be valid for natural proteins. Proper function in the original signaling context does not necessarily mean that they evolved to have large dynamic ranges or high catalytic activities. All properties might instead be tuned to the sweet spot of the host system so that testing linker variants can potentially optimize natural proteins, like the blue-light induced adenylate cyclase bPAC[90], for heterologous application.

It should be noted that using a starting construct of multiple protein modules, consecutive PATCHY reactions enable assembling receptor proteins with multiple sensors that integrate different signal inputs[72] or insert a module into existing protein architectures (Figure 21). Instead of randomizing the linker sequence completely, PATCHY generates all desired combinations of two parent linkers. Indeed, excessive sequence randomization would yield huge libraries with a

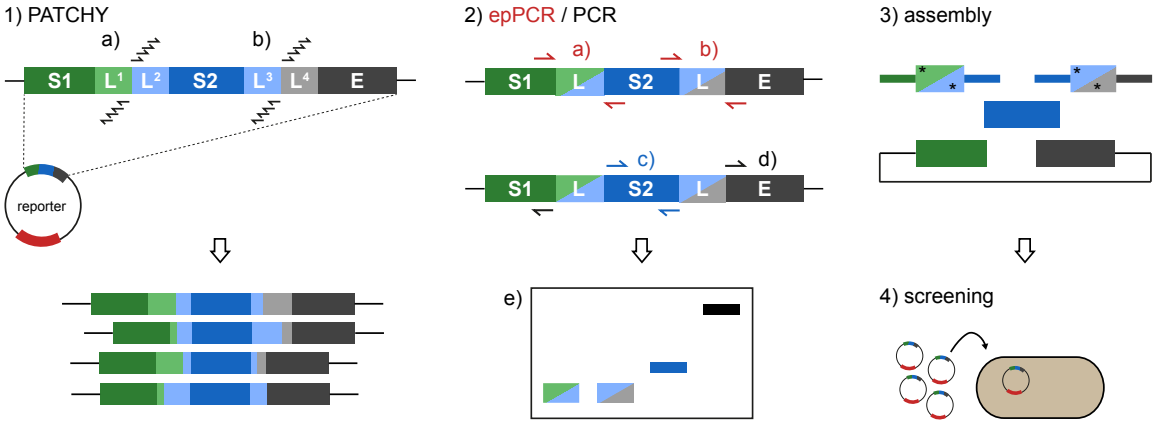

Figure 21: Proposed workflow for library-based chimera generation. Consecutive rounds of PATCHY generate multi-domain fusions with various linker combinations (**1**). Error-prone PCR of the linker regions (**2a, b**) adds mutations to tweak signaling across the linkers. A parallel PCR amplifies interjacent sensors (**2c**) and the plasmid backbone (**2d**). After purification via gel extraction (**2e**) all fragments are assembled using Gibson assembly or other restriction-free cloning strategies (**3**). Resulting fusion libraries are transformed e.g. into *E. coli* for functional testing (**4**).

vast majority of non-helical linkers and non-functional constructs. However, employing primers with few degenerate codons in the PATCHY PCR would introduce defined sequence variation. Especially at the fusion site this might repair eventual irregularities in the helix. Subsequent error-prone PCR of the linker or the whole protein can further tweak signaling between the modules[160]. A possible workflow includes one or multiple rounds of PATCHY (depending on the target design) using degenerate primers, followed by error-prone PCR and reassembly of the resulting insert with the plasmid backbone, e.g. using Gibson or cPec cloning[191] (Figure 21). Because all parts of the protocol involve PCR, no additional transformation steps are needed to amplify the amount of DNA.

Screening the libraries for light-regulated fusions depends on an assay that allows high-throughput testing. This study used bacterial assays, where histidine-kinase or adenylate-cyclase activity was coupled to the production of a red-fluorescent protein (3.3.1 and 5.2.1). Functional testing of eukaryotic or certain membrane proteins might require alternative test systems, e.g. yeast or eukaryotic cell culture. Also automated, multiwell-based purification and *in vitro* testing of proteins is conceivable. However, bacteria are most convenient due to their fast growth, easy cultivation and genetic manipulation. A major advantage of fluorescent reporters is that even visual inspection can quickly assess whether any functional constructs are present. Second, they allow screening via flow cytometry, which can quickly focus the library on single well-working constructs and speed up tool development.

6.3 Signal transduction in homodimeric photoreceptors

In modular, homodimeric photoreceptors two monomers often assemble in a parallel fashion with the helical linkers that connect sensors and effectors forming a central coiled-coil spine (3.2.1 and 3.3.1). Structural changes in the sensor propagate via these signaling helices and rearrange the tertiary or quaternary structure of the dimer and shifts the equilibrium of the effector modules between states of different biological activity[5]. Different photosensor and effector modules occur in nature in various combinations. It is reminiscent of Lego bricks, where modules are interchangeable, which suggests that signal transduction might follow common principles among these proteins. Obvious questions are: What are the signal mechanisms? Which particular residues in the protein and properties of the helical linker are critical? And how well are mechanisms conserved among certain groups of sensors, effectors or sensor-effector pairs? Hence, which modules are functionally interchangeable? Section 3.2 introduced several fundamentally different signaling mechanisms in sensory photoreceptors. Here, the discussion solely focuses on constitutively homodimeric proteins that undergo light-dependent changes of the tertiary or quaternary structure.

6.3.1 Signaling in YtvA-FixL fusions

The engineered light-regulated histidine kinase YF1 is a rare example where high-resolution structural data is available for the full-length homodimeric (photo-) receptor protein. Thus, it

represents a convenient model system for investigating the communication between the protein modules. In YF1 the *Bs*YtvA-LOV domain is fused to the *Bj*FixL-histidine kinase at the prominent DIT motif. Ergo this chimera only contains the linker from *Bj*FixL. In this study, the PATCHY protocol (4.4.2 and 5.3.1) was used to generate libraries of YtvA-FixL fusions with all possible combinations of the parental linkers (6.2.1). Evaluating the resulting light-regulated variants and their linker compositions gives insights into structural details and general principles that are key to signal transduction between LOV and kinase module. One of those crucial details appears to be the DIT motif mentioned above. Even though the library approach allowed variants lacking these residues partially or completely, they were preserved in all but one light-regulated YtvA-FixL fusions (5.3.2). The motif is highly conserved among LOV and other PAS domains and forms a salt bridge (between side chains of D125 and K96) and a hydrogen bond (between side chain of T127 and backbone amide of W103) to the LOV-domain core[63,78,192,193](Figure 22a and b). In a single construct (YF_L15) DIT is replaced by serine-arginine-leucine with the arginine residue most likely forming similar contacts to W103 and E105 of the domain core (Figure 22c). Data on the isolated YtvA-LOV domain indicated particularly low structural differences in this region between light and dark state so that it provides a rigid structure that couples the structural changes in the LOV-domain core to the linker helix[63,78]. An equivalent feature may exist in bacteriophytochromes, where the R-W-E/D motif forms a conserved salt bridge that couples the structural changes in the GAF loop to the linker helix[194]. This suggests pivots that feed the signal into the linker helix to be a common theme among various photoreceptor classes. The DIT motif is situated before the C-terminal linker of PAS domains and provides the only contact between domain core and C-terminal linker helix. In CBCR- and other GAF domains, however, the linker is situated at the dimer interface between the core domain and several residues form contacts to the GAF core so that signaling might be particularly susceptible to alterations of physical linker properties.

As natural photoreceptor proteins harbor linkers of various lengths, the total length of the linker and the absolute distance of sensor and effector domain seem to be secondary for signal transduction. Persistence lengths, that quantify the stiffness of a polymer, of roughly 100 nm and 150 nm for α helices or coiled coils[195] indicate that these structures behave as rigid rods on the molecular level[192]. In light-regulated YtvA-FixL fusions linkers vary between 13 to 59 residues, corresponding to a total distance between LOV-domain and kinase of approximately 20 to 90 Å. The fact that a single sensor-effector pair retains strict light-dependent switching despite such drastic changes in linker length, underlines the robustness of the signal mechanism over long distances.

Nevertheless, the linker length is still critical for sensor-effector communication and protein function. Light-repressed YtvA-FixL variants primarily occur at linkers comprising 7n+2 residues, light-induced variants at 7n+3 (5.3.2). Similar inversion of the signal response in YtvA-FixL fusions was observed before[78]. While previous inverted variants showed roughly 100-fold lower catalytic activity than YF1, bacterial assays indicate no such loss for the present variants. Still, the applied pDusk assay gives indirect information on kinase activity and further *in vitro* data have to confirm these results. Furthermore, the present study supports the notion that light regulation and inversion of the signal response occur systematically over broad range of total linker lengths.

These heptad periodicities stem from the fact that seven amino acids constitute two complete turns in the coiled-coil linker helices and preserve the orientation of the effector module relative to the sensor, as well as the relative orientation of the two effectors in the dimer (Figure 22d). The difference of one residue in the 7n+2 and 7n+3 types shifts the angle by approximately 103° and

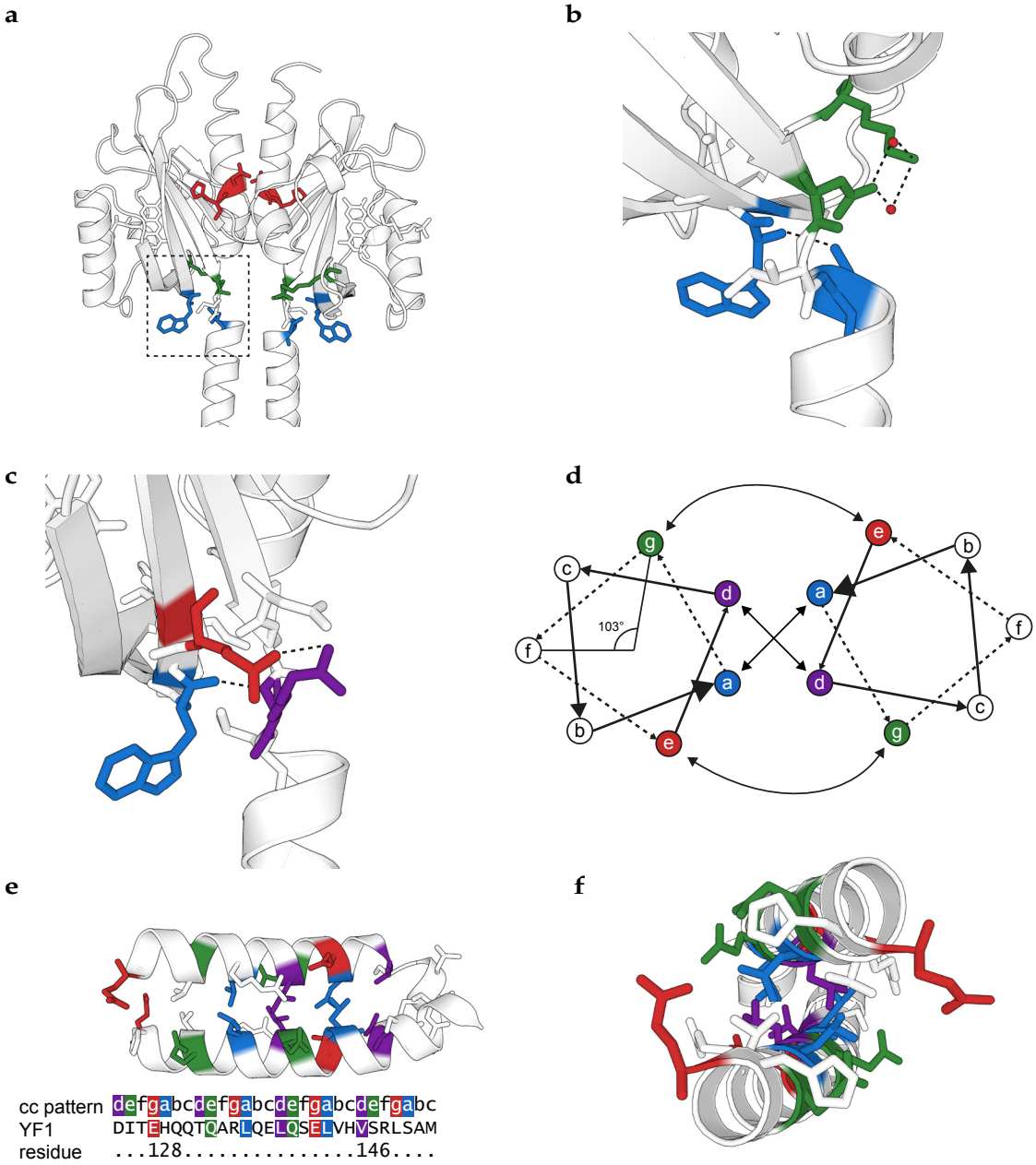

Figure 22: Signaling in YtvA-FixL fusions. (**a**) In the *Bs*YtvA-LOV domain mutations of D21 and H22 (red) invert the signal response. With polar contacts between K96 and D125 (green) as well as salt bridges between W103 and T127 (blue) the DIT motif provides a crucial pivot for propagating the structural change in the LOV core to the C-terminal signaling helices. (**b**) Polar contacts involving the DIT motif. Red spheres represent water as found in the original YF1 structure (PDB: 4GCZ) (**c**) Possible polar contacts in YF_L15, where SRL replaces DIT, involving R126 (purple), W103 (blue) and E105 (red). (**d**) Helix angles in a coiled coil change by 103° per residue. Hydrophobic interaction between non-polar residues at positions *a* and *d* as well as ionic interactions between charged residues at positions *e* and *g* stabilize the structure. (**e**) In the YF1 linker only a subset of residues (colored) complies with the regular coiled-coil pattern of non-polar and charged residues. (**f**) Top view of the YF1 linker showing stabilizing contacts within the coiled coil.

inverts the polarity of the light response. This suggests that the relative angular orientation of sensor and effector module determines the light response of the protein and the light signal itself elicits a structural change with a rotational component. Based on similar linker variants and structural data Möglich *et al.* developed a functional model that involves a torque motion around the interface of the LOV dimer and induces a supertwist in the coiled coil[67,78]. Still, there are exceptions to these heptad periodicities. One of the fusions that do stick to neither 7n+2 nor 7n+3 is YF_L15 (7n+1). This might be not that surprising as YF_L15 misses the DIT motif and contacts between LOV-domain core and linker helix may be completely different. Another construct is YF_L53 (7n+4), where a random mutation introduced a proline into the linker. The proline side chain interferes with the position of the helix backbone and potentially introduces a kink into the helix. Yet, changes in the linker do not necessarily have to be severe to change signal transduction between sensor and effector. Overall the number of identified, light-regulated YtvA-FixL fusions is remarkably low. Only 29 of 672 possible fusions and just a small fraction of fusions with linkers of 7n+2 or 7n+3 length were identified as light-switchable. Apart from minimal and maximal linker lengths multiple combinations of the parent YtvA and FixL linkers can implement each linker. However, usually only one or two light-regulated variants per linker length were isolated from the libraries. It is difficult to draw conclusions from the absence of specific constructs, because the reasons could be complex. One possibility is that the respective constructs have catalytic activities below the detection limit of the reporter assay; another is that the constructs were missing in the PATCHY libraries in the first place. Finally, single constructs could have been missed by chance. More detailed discussion of the library coverage (6.2.1) and the assays used for library screening (6.1.2) suggest that those reasons can account for the absence of single constructs, but not larger numbers. Thus, general trends like heptad periodicities or small fraction of light-regulated constructs are valid findings.

The amino acid sequence determines the molecular contacts in the linker helix and coiled-coil formation in the dimer. Generally, regular patterns of aliphatic and charged residues along the helix create hydrophobic interactions (between residues at positions *a* and *d*) and ionic interactions (between residues at positions *e* and *g*) that stabilize the coiled coil[196] (Figure 22d). These interactions make a coiled coil more rigid than a single α helix[195]. Furthermore, the rise of a coiled coil is 3.5 instead of 3.6 residues per turn so that the discussed angular orientations are slightly different compared to the α helix[197]. Hence, the exact mix of linker parts that adopt a coiled-coil structure or not could influence the functionality of the respective fusion protein. The coiled-coil linker of YF1/FixL does not fully comply with the typical regular pattern of hydrophobic and charged residues. Bioinformatic predictions[198] clearly suggest a coiled-coil formation between residues E128 and V146 (data not shown). Mapping the present contacts to the crystal structure however shows that not all *a* and *d* positions harbor hydrophobic or *e* and *g* positions charged side chains (Figure 22e, f). Deviation of the linker from an ideal coiled coil might represent a fine-tuning of linker composition to signaling in the particular protein context, e.g. by providing specific physical properties of the linker or inducing asymmetries in the dimer[67]. This fine-tuning might indeed pose the major complication of rewiring photoreceptor modules.
The YF_L58 T148A and YF_L59 A148L variants (5.3.2) introduce mutations that are subtle enough to conserve coiled-coil probabilities compared to YF_L58 and YF_L59 wild types (predictions

using COILS[198], data not shown). The fact that they still impair light regulation in the protein underlines the fragility of signal propagation through the linker and could account for the missing constructs that were expected to be light-regulated based on their linker length.

Sequence alignments show that natural PAS-HisK proteins also maintain heptad periodicities of linker lengths[78] (5.4). Two prominent populations occur at 7n+0 and 7n+2, with FixL belonging to the latter class. The majority of light-repressed YtvA-FixL fusions obey this 7n+2 pattern. In contrast, linker lengths of 7n+3, as in most light-induced variants, are completely missing in natural PAS-HisK proteins. A significant 7n+1 population of potential inverted variants of the 7n+0 class is absent as well, so that this inverter principle seems to be artificial. However, polarity inversion by addition of one linker residue might be transferable to a broad range of histidine kinases, at least to the ones that obey the 7n+2 pattern. The rationally designed YtvA-FixL fusion YF1 also demonstrated other ways to invert the polarity of the light response. Mutations at the interface between beta-sheet and A' helix of the LOV domain (residues D21 and H22, Figure 22a) render the light-repressed YF1 light-inducible[67,160]. Here as well, further experiments have to show whether this principle is applicable to other proteins. Whereas these mutations could provide a general strategy for polarity inversion in LOV- and PAS domains, the linker variations might work in various sensor-effector combinations. Consequently, it might be a widely applicable strategy to engineer receptors of complementary signal sensitivity.
The following sections deal with the signaling transduction in other homodimeric signal receptors (6.3.2) and discuss further experiments for testing conservation of specific mechanisms and transferability of functional principles among different them (6.3.3).

6.3.2 Generality of signaling mechanisms

Systematic rewiring of sensor and effector modules is one approach to test for common signaling principles. Analogous to the YtvA-FixL fusions, experiments in section 5.3.3 ought to investigate the compatibility of the *Bs*YtvA-LOV domain with other effector modules, i.e. the adenylate cyclases from *An*CyaB1 and bPAC. The YtvA-CyaB1 approach majorly failed due to technical problems. Robust background activity of most fusions as well as the isolated cyclase module hampered the identification of light-switchable constructs. Fusions of *Bs*YtvA and bPAC showed low, but significant light regulation (5.3.3), which suggests that the general signaling principles in both proteins, might be similar to some extent. Nevertheless, the YtvA-LOV domain might be not suited for stringent switching of the bPAC cyclase. Current data stem from bacterial reporter assays that indirectly reflect catalytic activities, but are bolstered by *in-vitro* measurements that confirmed induction factors of selected YtvA-bPAC fusions (5.3.3). However, due to the small number of light-regulated fusions and their extremely low dynamic range more detailed analysis is not useful at this point. Instead more information on general compatibility will require rewiring of *Bs*YtvA-LOV to different cyclases. The closely related mPAC features an N-terminal LOV domain and was reported to show blue-light regulated adenylate-cyclase activity[199]. Yet, with respect to the diverse light-induced structural changes in LOV domains (3.1.1 and 3.2), the signal transduction might be very well different compared to the YtvA-LOV domain. Generating

and testing YtvA-mPAC chimeras with PATCHY and the pCyclR reporter would allow further elucidation of this issue.

Complementary to the recombination of protein modules, alignment of natural proteins exposes patterns of linker lengths or conserved sequence motifs. The YtvA-FixL fusions discussed above (6.3.1), as well as the majority of natural PAS-HisK proteins, feature linker lengths complying with heptad periodicities. Corresponding analyses of other sensor-effector pairs showed that similar patterns occur in PAS-GGDEF[192], PAS-CYCc, GAF-HisK, GAF-GGDEF and GAF-CYCc proteins (5.4) and between multiple sensors[73]. The analysis neglects the possibility that some of the proteins are non-functional or do not belong to the class of homodimeric receptor proteins, but reveals general trends. As discussed before (6.3.1) linker lengths differing by seven residues conserve the relative orientation of the connected modules. Within a population, especially the proteins that feature these periodicities might have similar signaling mechanisms so that module swapping could be feasible among them.
Analogous comparison between different sensor-effector pairs is hampered by the fact that absolute linker lengths are not comparable, because the linker boundaries are arbitrary for each module combination. Instead conserved functional principles like linker-length dependent polarity inversion could indicate functional interchangeability. Whereas light-repressed and light-induced YtvA-FixL fusions differ by one residue, a study by Winkler *et al.* reports the signal inversion of an HAMP-coupled adenylate cyclase by adding four residues to the linker helix[181]. An engineered bacteriophytochrome-adenylate cyclase shows signal-dependency of both polarities at linker length differing by three or four residues[157]. Finally, linker variation in an engineered bacteriophytochrome-phosphodiesterase did not yield any additional light-regulated fusion[158]. Yet, all these studies tested just few fusion variants and systematic linker variation will probably identify rules of more general validity.
Multiple prolines, found in CYCc and PDE linkers, could break the regular structure of the α helix and induce irregularities in the coiled-coil linker. Implications for the signaling mechanism will need further investigation. However, prolines might create kinked helices and protein geometries specific for cyclases and phosphodiesterases.

High-resolution structural data form another pillar for investigating intramolecular signaling mechanisms. The eminent lack of structural data on full-length proteins, as well as missing structures of both dark and lit states, impedes deeper insights. The torque model (Figure 23a) derived from structural and functional data on YF1 is in line with other studies on histidine kinases that propose conformational changes with helix rotation as the principal component[200–204]. Associated structural rearrangements could alter the interaction and mobility of the DHp and CA domains of the kinase[67,205] and switch the effector between net kinase and phosphatase activity. Transmembrane chemoreceptors featuring HAMP-coupled histidine kinases might also undergo piston-type structural changes[206]. Recent data on *Deinococcus radiodurans* bacteriophytochrome (*Dr*BPhy) reveal conformational changes in the photosensor that regulates a histidine kinase in response to red and far-red light. Crystal structures and solution scattering data of the isolated PAS-GAF-PHY sensor module point toward an opening of the homodimer in a scissor-like motion in the red-light adapted Pfr state[68,207,208] (Figure 23b). A reorientation around the swivel

point of the PAS-GAF interface in the dimer leads to a rotation and outward movement of the PHY domains and the attached linker helices[209]. This model shares characteristics with structural changes in the LOV dimer, where rotation around the interface of β sheets and A'α helices elicits rotation and scissor motion of the Jα linker helices upon illumination[63,67]. The more pronounced structural change in the bacteriophytochrome could stem from the fact that it was crystalized in dark and light state, whereas in the case of YtvA-LOV dark-adapted crystals were subsequently illuminated with light. Hence, crystal packing might have precluded major structural rearrangements. Another reason could be the domain architecture of the photosensor module. The PAS-GAF-PHY module provides a much larger lever than the single LOV domain. Indeed, different sensors might amplify small structural changes around the chromophore in different fashions, which could limit the interchangeability of protein modules. Furthermore, electron microscopy of the full-length *Dr*BPhy suggests a heterogeneous Pfr state constituting an equilibrium of two conformations[209]. In one of them the PHY domains are moved apart so that the relative orientation of the attached histidine-kinase modules is most likely altered, whereas in the second one the kinases seem to be largely separated (Figure 23b). It is unknown if this mechanism is unique to phytochromes, bacteriophytochromes or phytochrome-histidine kinases or whether it has functional relevance. Similar dissociation of the effector modules could also happen in other proteins and does not contradict the mechanistic aspects for histidine kinases discussed above. A crystal structure and X-ray scattering data of the light-adapted state of YF1 could contribute to

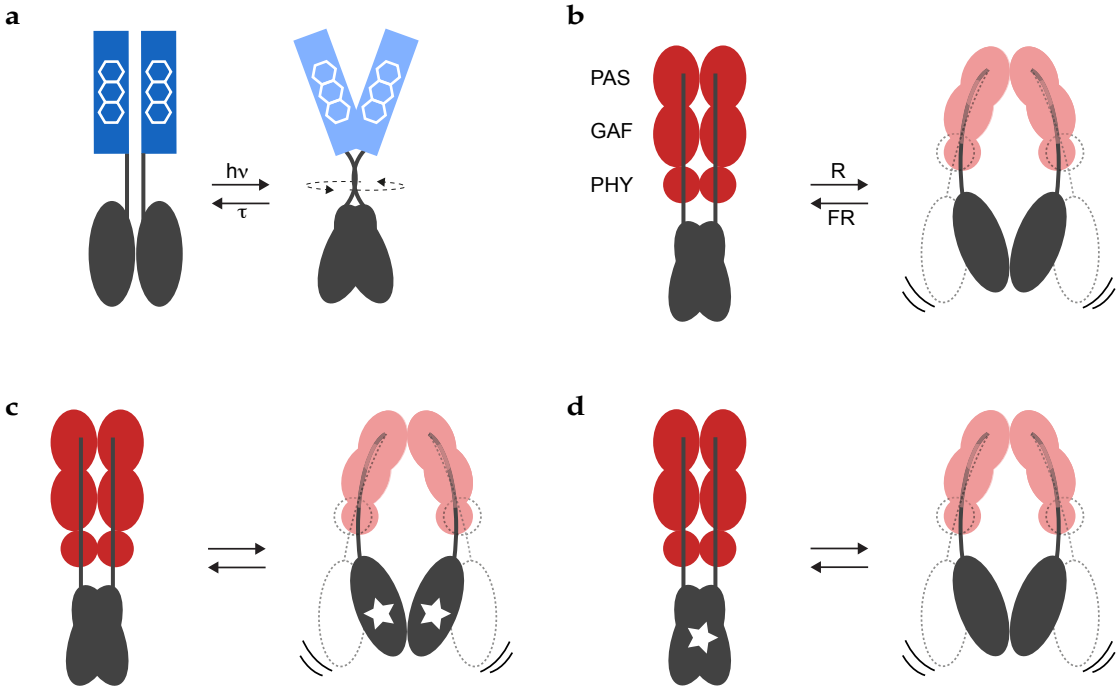

Figure 23: Signaling in homodimeric receptor proteins. (**a**) A light-induced torque motion in the YF1 dimer is proposed for switching the effector between net kinase and phosphatase activity. (**b**) In phytochromes red light induces reorientation around the PAS-GAF dimer interface that repositions the PHY domain and the effector module. Electron microscopy of *Dr*BPhy indicates histidine-kinase effectors to dissociate completely (adapted from Burgie *et al.*, 2014). (**c**) Phosphodiesterases adopt an R state where the effector monomers block each other's catalytic sites. In the relaxed state the effectors dissociate and become active. (**d**) In adenylate/guanylate cyclases the constrained state inhibits the correct assembly of both effector monomers to form a catalytic site at the dimer interface. Stars indicate catalytic activity.

further investigation. The engineered red-light induced phosphodiesterase LAPD proved that *Dr*BPhy is able to switch the activity of diverse effector modules[158]. Furthermore, the analogous bacteriophytochrome from *Rhodobacter sphaeroides* was shown to control an adenylate cyclase[157] in a red-light dependent manner.

Möglich *et al.* described a common light-induced torque movement[33,67,78] that Schultz and Natarajan incorporated into a general light-induced 'regulated unfolding' of the linker that could involve rotation and partial unwinding of the coiled coil[77]. Receptor proteins exist in equilibrium between a constrained state of low output activity and a relaxed state of enhanced activity[210], where the effector modules in the dimer can orient according to their intrinsic affinities[77]. The structural change in the linker shifts the equilibrium between these states and alters the net activity of the protein. Conformational changes in the respective effector module might well be different, so that each effector interprets the signal differently. Whereas PDEs block each other in the dimer and disassemble in a signal-dependent manner (Figure 23c), adenylate cyclases are inactive as monomers and form a catalytic site at the dimer interface[181,211] (Figure 23d). In histidine kinases this transition might switch the effector between net kinase and net phosphatase activity. Thereby, the linker properties, i.e. its length, define the configuration and activity of the basal state. In summary, structural data on phosphodiesterases[79], adenylate cyclases[212–214] and c-di-GMP cyclases[215] do not contradict shared signaling mechanisms involving rotational motion suggested for histidine kinases. However, the linker sequence might tune signaling in specific sensor-effector pairs, e.g. by deviation from ideal coiled-coil patterns (6.3.1) or proline-induced kinks (5.4 and above discussion), and hamper module swapping between different proteins.

6.3.3 Further investigation of intramolecular signal transduction

Key challenges for a deeper understanding of the signal transduction between sensor and effector modules involve the structural changes during signaling, their general validity among different sensor-effector pairs and the connection between linker sequence and protein function.
On the one hand, extensive structural studies on natural photoreceptor proteins report on the general arrangement of sensor and effector modules and their signal-induced dynamics. On the other hand, systematic rewiring of sensor and effector modules allows deducing more information on their compatibility. The previous sections indicated that the connection between protein sequence and function is not understood well enough to deduce signaling mechanisms or module compatibility. For example among LOV proteins, light triggers highly diverse structural changes (3.1.1 and 3.2). Still, it is not possible to distinguish mechanisms based on the sensor amino-acid sequence or predict how structural changes in the sensor translate to a different effector. In order to understand how specific structural cues e.g. the linker sequence determines protein functionality, it is not sufficient to investigate signaling in natural photoreceptor proteins. Instead it is necessary to look at how the exact same sensors or effectors work with different partners. As both proteins feature the same effector module, in particular structural data of the CF chimera (5.1.2) may serve for a direct comparison with YF1 (3.3.1). The same applies to YF1 and the original YtvA, which share the same sensor.

Moreover, studies on linker variants with different signal polarities of a particular sensor-effector pair can help to understand how the linker sequence and single-residue changes influence the protein function. Static and dynamic structural data of light-repressed and -induced YtvA-FixL linker variants would elucidate, whether the 'light structure' of light-induced variants resemble dark-state YF1 (and the other way around). And how structures of inverted YtvA-FixL variants from linker variation (5.3.2) or mutation (D21 and H22 mutations, see 6.3.1) in the LOV domain compare to each other. Next steps to test the transferability of both inverter principles should include systematic linker variation (e.g. using PATCHY) in other FixL chimeras like CF and introduction of D21/H22 mutations into new LOV-FixL fusions. These experiments may employ the pDusk assay for functional screening. Testing signal inversion in proteins harboring other effector modules will need new reporter assays with sufficient throughput, e.g. the adenylate-cyclase reporter described in section 5.2.1.

6.4 Toward novel molecular tools

Complementary to designed tools, genome screening revealed widely-used optogenetic actuators like the strongly red-shifted channelrhodopsin Chrimson[216] or the blue-light induced adenylate cyclase bPAC[90]. As these proteins were optimized in the course of evolution, they often outperform engineered constructs. Therefore, the quest for optogenetic tools or other molecular switches should start out by screening databases for natural proteins featuring the desired traits. Yet, a precondition is sufficient knowledge about the protein to deduce specific functional properties from the amino-acid sequence. Protein engineering becomes especially relevant when no natural proteins are available to control a certain cellular process or for optimizing natural proteins to heterologous applications. Also, effector functions vary in their specificity. While functions like ion conductance across the plasma membrane or production of a second messenger represent non-specific signals, functions like phosphorylation, protein-protein interaction or DNA-binding are highly specific for a certain target. Hence, if signaling paths involve for example the phosphorylation by a kinase, it is insufficient to take any light-regulated kinase, but demands rendering the specific kinase light sensitive.

Engineering chimeric photoreceptors often employs associating types (3.2). Mediating light-dependent recruitment of an effector to a target site or association of a split protein represent comparatively straightforward design strategies. As in this case the linker is usually unstructured and primarily connects sensor and effector modules to maintain spatial proximity. Thus, fusing proteins at these sites is usually less critical than in linkers of homodimeric photoreceptors discussed in the present study. The light-dependent unfolding of the C-terminal Jα helix in *As*LOV2 embodies another versatile building brick for reversible blocking of effectors and epitopes (3.2). Most studies use the original linker with few modifications of its lengths and sequence to optimize light regulation. Additionally, several mutations are available that modulate the dynamic range of the light response[108,109].

Recombination of homologous sensor and effector modules facilitates light control of homodimeric proteins. As seen in the present study, signaling within and between respective protein modules is still obscure and their interchangeability rather unpredictable. Hence, successful design might require testing several sensor or effector modules. In addition, the physical nature of the linker crucially affects the signal propagation between the modules and the functionality of the protein (5.3.2). Consequently, numerous linker designs might be tested to yield a light-regulated chimera. The PATCHY approach presented in this study can help with the latter issue and significantly accelerate the engineering of chimeric proteins. Still, the strategy relies on efficient test setups that allow functional screening with high throughput. Based on the insights from the present study, the following sections discuss approaches for revisiting the design of chimeric CBCR photoreceptors and engineering novel molecular tools.

6.4.1 Cyanobacteriochromes

Sections 5.1 and 5.2.2 reported the rational design of chimeric cyanobacteriochrome proteins. As pointed out before, their bistability and spectral diversity are highly favorable traits for optogenetic applications. While the latter facilitates parallel application of tools with orthogonal light sensitivities, bistable photocycles allow spatially defined activation and deactivation of the photoreceptor. Resembling a STED (stimulated emission depletion) microscopy setup[217], superposition of activating and deactivating light beams could spatially confine the activity of a CBCR-coupled effector with subcellular precission[5] (Figure 24c). A major issue for heterologous application however is the bilin chromophore supply. While biliverdin used by bacterial phytochromes is available in a wide range of tissues and ubiquitous in mammalian tissues[94,158,157,218,219], PCB has to be exogenously delivered to most systems other than plants. Although, PCB synthesis was established in mammalian cells, it still complicates the application of CBCRs for optogenetic purposes. The recent discovery of the cyanobacteriochrome AM1_1557 from *Acaryochloris marina* may pave the way to solving this problem[220]. Upon PCB incorporation it shows a typical R/G photocycle. However, it binds biliverdin with similar affinity leading to a far-red/orange photocycle. The respective study suggests that a leucine residue (L337) primarily contributes to biliverdin binding[220]. Alignments with other CBCRs show that this position usually harbors a conserved asparagine or histidine residue in R/G CBCRs and a cysteine in other subtypes (Figure 24a). Thus, corresponding mutations might facilitate biliverdin incorporation in a wide range of cyanobacteriochromes and make them far more attractive for heterologous application. As the authors of the respective study point out, AM1_1557 retains residual BV binding upon mutating the leucine to asparagine and deduce that other residues might also contribute. Thus, random mutagenesis of CBCRs harboring the N-to-L mutation supposedly provides the most promising approach. Screening mutation libraries might be possible by flow cytometry harnessing the CBCR fluorescence that significantly differs for PCB or BV binding variants[220]. Using a similar strategy can also enable optimizing CBCRs for application as bistable fluorescent reporters[220]. The CcaS-FixL chimera CF (5.1.2) constitutes an apparent target for this mutation, because it readily reports functional signal transduction of the mutants.

Successful rewiring of a CBCR sensor to the FixL histidine kinase in light-regulated CF chimera represents the only design that comprised a CBCR sensor in its original protein context, ergo features the complete *Sy*CcaS sensor including its two PAS domains at the N-terminus of the protein. Novel designs should therefore focus on either single CBCRs that reside directly N-terminal of an effector module or use complete sensor modules composed of different sensors. *Sy*CcaS is perfect in that context, because is comprises just one photosensitive domain. Other CBCRs usually occur in tandem with other variants of different spectral sensitivities (3.1.2) and would provide complex inputs to the effector. However, *Np*F2854 features a sensor module composed of four R/G CBCR sensors and would allow a specific response to red and green light (Figure 24b). Data mining also revealed single CBCRs that are directly coupled to a C-terminal histidine kinase (Figure 24b). While the corresponding response regulators of the kinases are unknown so far, there is a good chance that these CBCRs are functionally compatible to kinases of known specificity, like FixL, or serve as building bricks for light-regulation of other effectors.

A different approach for rewiring CBCRs to new effectors might involve just changing the GAF-domain core. This strategy either applies to render already GAF-coupled effectors light-sensitive or generate spectrally orthogonal variants of a functional CBCR chimera. Starting out from CF the *Sy*CcaS-GAF domain could be exchanged for CBCR sensors with orthogonal light sensitivities by replacing the core β sheet and parts of the N- and C-terminal helices that make contacts to the

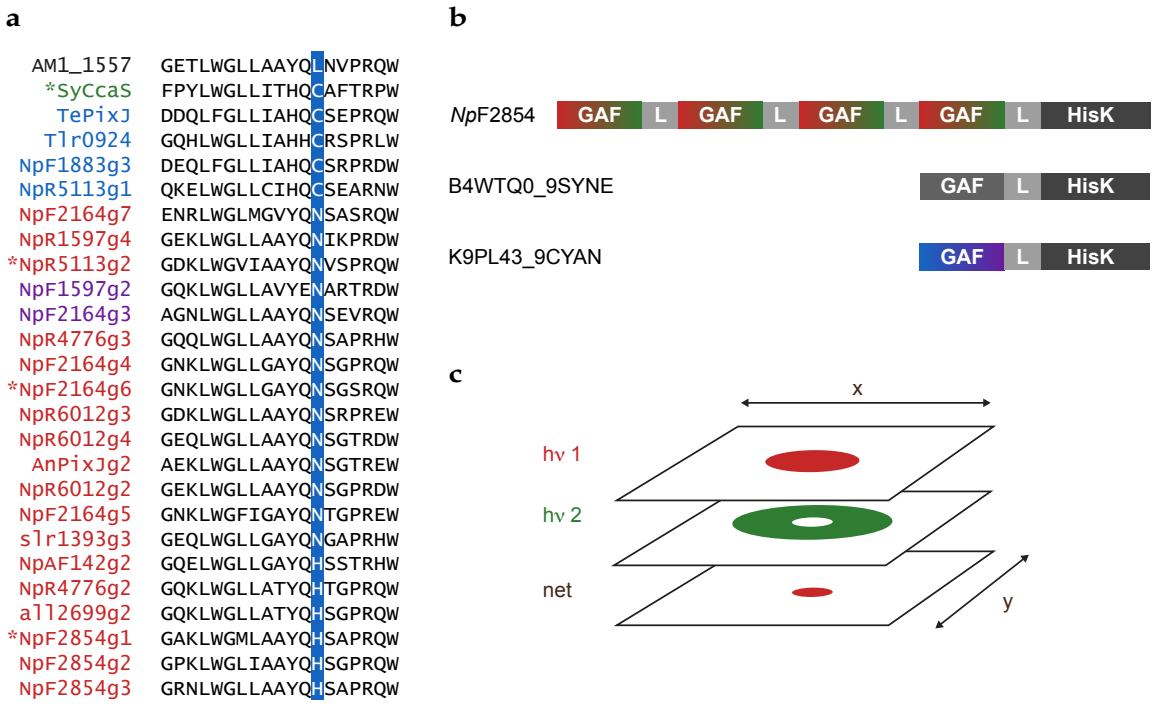

Figure 24: Targets for CBCR engineering. (**a**) Sequence alignment showing a leucine residue in the biliverdin-binding AM1_1557 at a position where other CBCRs feature conserved asparagine or cysteine residues. CBCR occurring in the present study are marked with ' * '. (**b**) CBCRs that could provide building bricks for further chimera generation. (**c**) Superposed light beams of activating (hv1) and deactivating (hv2) wavelengths might facilitate stimulation of CBCRs and other photochromic photoreceptor proteins with supreme spatial resolution (Adapted from Ziegler and Möglich 2015).

core. Meanwhile, parts of N- and C-terminus that protrude from the sensor would be conserved (see Figure 3b for CBCR structure). Following this strategy might also allow incorporating CBCR sensors that originally occur in the middle of a protein (e.g. *Np*R5113_GAF2 from section 5.1). At this point it is hard to predict the structure and sequence cues that determine whether a sensor works in a certain protein context. Especially, the N-terminus is usually unstructured and not resolved in crystal structures. Therefore, selecting the exact boundaries for the insertion of a new sensor remains difficult. Using a similar approach described in 6.2.2, two consecutive PATCHY reactions are able to vary fusion at both sites and efficiently test for functional constructs.
Screening databases for promising protein architectures will readily complement these design strategies.

6.4.2 Other receptor proteins

Bacterial phytochromes were recently applied for generating homodimeric photoreceptor chimeras[157,158,221]. Like CBCRs they have a bistable photocycle, but the absorption spectra of the two light absorbing states Pr and Pfr overlap significantly and could hamper specific population of either photostate. Recently found algal phytochromes show a vast spectral diversity comparable to CBCRs[6] and may complement present phytochrome chimeras. Yet, they need PCB and are therefore mainly attractive for plant, prokaryotic and selected mammalian systems where the chromophore is either endogenous or can be readily delivered[169,170,222].
A novel rhodopsin-coupled guanylate cyclase was reported to allow optical control of cellular cGMP levels and exceled with extremely low dark activity[223]. Similar enzyme-coupled rhodopsins also occur with other effectors like histidine kinases[224]. Rewiring these rhodopsin sensors to new effectors was not reported so far. As the linker between sensor and effector most likely transmits a structural change like in homodimeric proteins in the present study, engineering might require testing numerous linker variants, which could be created with PATCHY. As many rhodopsins, e.g. channelrhodopsins, poorly express in prokaryotes establishing high-throughput screen poses a major problem. However, fluorescence-based reporter assays allowing selection via flow-cytometry can be also set up in mammalian cells and already exist e.g. for adenylate/guanylate cyclases and phosphodiesterases[223,225]. Viral vectors could deliver fusion libraries. For example, application of adenovirus was shown to yield clonality in human embryonic kidney cells transfected with antibody libraries[226]. Therefore, engineering chimeric enzyme-coupled rhodopsins could be well within reach.

Beyond controlling signaling pathways, engineered receptor proteins may also report on cellular signals. The PAS domain family, that includes the blue-light sensitive LOV domains, comprises modules that specifically bind diverse metabolites or divalent cations like calcium[166,227]. The same is true for GAF domains that were shown to sense cyclic nucleotides[79,178,228]. However, these sensors were not systematically studied so far and for most of them the ligand remains unknown. Still, rewiring these sensor modules to enzymatic outputs or gene transcription has the potential for establishing *in-vivo* reporters for diverse metabolic processes. Inspired by the diversity of natural PAS domains, mutational studies should furthermore aim to change specificities toward

novel ligands. In this context, engineered or natural metabolite-sensing PAS domains that exhibit helix unfolding similar to *As*LOV2 could significantly simplify design approaches.

7 Bibliography

(1) Christie, J. M., Reymond, P., Powell, G. K., Bernasconi, P., Raibekas, A. A., Liscum, E., and Briggs, W. R. (1998) Arabidopsis NPH1: a flavoprotein with the properties of a photoreceptor for phototropism. *Science 282*, 1698–1701.

(2) Nagel, G., Ollig, D., Fuhrmann, M., Kateriya, S., Musti, A. M., Bamberg, E., and Hegemann, P. (2002) Channelrhodopsin-1: a light-gated proton channel in green algae. *Science 296*, 2395–2398.

(3) Gehring, W. J. (2005) New Perspectives on Eye Development and the Evolution of Eyes and Photoreceptors. *J. Hered. 96*, 171–184.

(4) Möglich, A., Yang, X., Ayers, R. A., and Moffat, K. (2010) Structure and function of plant photoreceptors. *Annu. Rev. Plant Biol. 61*, 21–47.

(5) Ziegler, T., and Möglich, A. (2015) Photoreceptor engineering. *Front. Mol. Biosci. 2*.

(6) Rockwell, N. C., Duanmu, D., Martin, S. S., Bachy, C., Price, D. C., Bhattacharya, D., Worden, A. Z., and Lagarias, J. C. (2014) Eukaryotic algal phytochromes span the visible spectrum. *Proc. Natl. Acad. Sci. 111*, 3871–3876.

(7) Herrou, J., and Crosson, S. (2011) Function, structure and mechanism of bacterial photosensory LOV proteins. *Nat. Rev. Microbiol. 9*, 713–723.

(8) Taylor, B. L., and Zhulin, I. B. (1999) PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev. 63*, 479–506.

(9) Möglich, A., Ayers, R. A., and Moffat, K. (2009) Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure 17*, 1282–1294.

(10) Möglich, A., and Moffat, K. (2007) Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. *J. Mol. Biol. 373*, 112–126.

(11) Möglich, A., Yang, X., Ayers, R. A., and Moffat, K. (2010) Structure and function of plant photoreceptors. *Annu. Rev. Plant Biol. 61*, 21–47.

(12) Schleicher, E., Kowalczyk, R. M., Kay, C. W. M., Hegemann, P., Bacher, A., Fischer, M., Bittl, R., Richter, G., and Weber, S. (2004) On the reaction mechanism of adduct formation in LOV domains of the plant blue-light receptor phototropin. *J. Am. Chem. Soc. 126*, 11067–11076.

(13) Salomon, M., Christie, J. M., Knieb, E., Lempert, U., and Briggs, W. R. (2000) Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry (Mosc.) 39*, 9401–9410.

(14) Crosson, S., Rajagopal, S., and Moffat, K. (2003) The LOV domain family: photoresponsive signaling modules coupled to diverse output domains. *Biochemistry (Mosc.) 42*, 2–10.

(15) Lederer, F. (1992) Extreme pKa displacements at the active sites of FMN-dependent α-hydroxy acid-oxidizing enzymes. *Protein Sci. 1*, 540–548.

(16) Crosson, S., and Moffat, K. (2002) Photoexcited Structure of a Plant Photoreceptor Domain Reveals a Light-Driven Molecular Switch. *Plant Cell 14*, 1067–1075.

(17) Zoltowski, B. D., Schwerdtfeger, C., Widom, J., Loros, J. J., Bilwes, A. M., Dunlap, J. C., and Crane, B. R. (2007) Conformational switching in the fungal light sensor Vivid. *Science 316*, 1054–1057.

(18) Vaidya, A. T., Chen, C.-H., Dunlap, J. C., Loros, J. J., and Crane, B. R. (2011) Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription. *Sci. Signal. 4*, ra50.

(19) Yee, E. F., Diensthuber, R. P., Vaidya, A. T., Borbat, P. P., Engelhard, C., Freed, J. H., Bittl, R., Möglich, A., and Crane, B. R. (2015) Signal transduction in light-oxygen-voltage receptors lacking the adduct-forming cysteine residue. *Nat. Commun. 6*, 10079.

(20) Kasahara, M., Swartz, T. E., Olney, M. A., Onodera, A., Mochizuki, N., Fukuzawa, H., Asamizu, E., Tabata, S., Kanegae, H., Takano, M., Christie, J. M., Nagatani, A., and Briggs, W. R. (2002) Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from Arabidopsis, rice, and Chlamydomonas reinhardtii. *Plant Physiol. 129*, 762–773.

(21) Purcell, E. B., Siegal-Gaskins, D., Rawling, D. C., Fiebig, A., and Crosson, S. (2007) A photosensory two-component system regulates bacterial cell attachment. *Proc. Natl. Acad. Sci. U. S. A. 104*, 18241–18246.

(22) Christie, J. M., Swartz, T. E., Bogomolni, R. A., and Briggs, W. R. (2002) Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function. *Plant J. 32*, 205–219.

(23) Kennis, J. T. M., van Stokkum, I. H. M., Crosson, S., Gauden, M., Moffat, K., and van Grondelle, R. (2004) The LOV2 domain of phototropin: a reversible photochromic switch. *J. Am. Chem. Soc. 126*, 4512–4513.

(24) Harper, S. M., Neil, L. C., and Gardner, K. H. (2003) Structural basis of a phototropin light switch. *Science 301*, 1541–1544.

(25) Avila-Pérez, M., Hellingwerf, K. J., and Kort, R. (2006) Blue light activates the sigmaB-dependent stress response of Bacillus subtilis via YtvA. *J. Bacteriol. 188*, 6411–6414.

(26) Jurk, M., Dorn, M., Kikhney, A., Svergun, D., Gärtner, W., and Schmieder, P. (2010) The switch that does not flip: the blue-light receptor YtvA from Bacillus subtilis adopts an elongated dimer conformation independent of the activation state as revealed by a combined AUC and SAXS study. *J. Mol. Biol. 403*, 78–87.

(27) Losi, A., Polverini, E., Quest, B., and Gärtner, W. (2002) First evidence for phototropin-related blue-light receptors in prokaryotes. *Biophys. J. 82*, 2627–2634.

(28) Losi, A., Quest, B., and Gärtner, W. (2003) Listening to the blue: the time-resolved thermodynamics of the bacterial blue-light receptor YtvA and its isolated LOV domain. *Photochem. Photobiol. Sci. 2*, 759–766.

(29) Losi, A., Gärtner, W., Raffelberg, S., Cella Zanacchi, F., Bianchini, P., Diaspro, A., Mandalari, C., Abbruzzetti, S., and Viappiani, C. (2013) A photochromic bacterial photoreceptor with potential for super-resolution microscopy. *Photochem. Photobiol. Sci. 12*, 231.

(30) Losi, A., Ternelli, E., and Gärtner, W. (2004) Tryptophan Fluorescence in the Bacillus subtilis Phototropin-related Protein YtvA as a Marker of Interdomain Interaction¶. *Photochem. Photobiol. 80*, 150–153.

(31) Losi, A., Ghiraldelli, E., Jansen, S., and Gärtner, W. (2005) Mutational effects on protein structural changes and interdomain interactions in the blue-light sensing LOV protein YtvA. *Photochem. Photobiol. 81*, 1145–1152.

(32) Möglich, A., Ayers, R. A., and Moffat, K. (2009) Design and signaling mechanism of light-regulated histidine kinases. *J. Mol. Biol. 385*, 1433–1444.

(33) Möglich, A., and Moffat, K. (2010) Engineered photoreceptors as novel optogenetic tools. *Photochem. Photobiol. Sci. 9*, 1286–1300.

(34) Ikeuchi, M., and Ishizuka, T. (2008) Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. *Photochem. Photobiol. Sci. 7*, 1159.

(35) Hirose, Y., Shimada, T., Narikawa, R., Katayama, M., and Ikeuchi, M. (2008) Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phycobilisome linker protein. *Proc. Natl. Acad. Sci. U. S. A. 105*, 9528–9533.

(36) Rockwell, N. C., and Lagarias, J. C. (2010) A Brief History of Phytochromes. *ChemPhysChem 11*, 1172–1180.

(37) Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature 438*, 325–331.

(38) Song, C., Psakis, G., Lang, C., Mailliet, J., Gärtner, W., Hughes, J., and Matysik, J. (2011) Two ground state isoforms and a chromophore D-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome. *Proc. Natl. Acad. Sci. 108*, 3842–3847.

(39) Yang, X., Kuk, J., and Moffat, K. (2008) Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: photoconversion and signal transduction. *Proc. Natl. Acad. Sci. 105*, 14715–14720.

(40) Rockwell, N. C., Martin, S. S., Feoktistova, K., and Lagarias, J. C. (2011) Diverse two-cysteine photocycles in phytochromes and cyanobacteriochromes. *Proc. Natl. Acad. Sci. U. S. A. 108*, 11854–11859.

(41) Hirose, Y., Rockwell, N. C., Nishiyama, K., Narikawa, R., Ukaji, Y., Inomata, K., Lagarias, J. C., and Ikeuchi, M. (2013) Green/red cyanobacteriochromes regulate complementary chromatic acclimation via a protochromic photocycle. *Proc. Natl. Acad. Sci. 110*, 4974–4979.

(42) Narikawa, R., Ishizuka, T., Muraki, N., Shiba, T., Kurisu, G., and Ikeuchi, M. (2013) Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci. U. S. A. 110*, 918–923.

(43) Rockwell, N. C., Martin, S. S., Gulevich, A. G., and Lagarias, J. C. (2014) Conserved Phenylalanine Residues Are Required for Blue-Shifting of Cyanobacteriochrome Photoproducts. *Biochemistry (Mosc.) 53*, 3118–3130.

(44) Rockwell, N. C., Martin, S. S., Gan, F., Bryant, D. A., and Lagarias, J. C. (2015) NpR3784 is the prototype for a distinctive group of red/green cyanobacteriochromes using alternative Phe residues for photoproduct tuning. *Photochem Photobiol Sci 14*, 258–269.

(45) Ishizuka, T., Narikawa, R., Kohchi, T., Katayama, M., and Ikeuchi, M. (2007) Cyanobacteriochrome TePixJ of Thermosynechococcus elongatus harbors phycoviolobilin as a chromophore. *Plant Cell Physiol. 48*, 1385–1390.

(46) Ishizuka, T., Kamiya, A., Suzuki, H., Narikawa, R., Noguchi, T., Kohchi, T., Inomata, K., and Ikeuchi, M. (2011) The cyanobacteriochrome, TePixJ, isomerizes its own chromophore by converting phycocyanobilin to phycoviolobilin. *Biochemistry (Mosc.) 50*, 953–961.

(47) Enomoto, G., Hirose, Y., Narikawa, R., and Ikeuchi, M. (2012) Thiol-Based Photocycle of the Blue and Teal Light-Sensing Cyanobacteriochrome Tlr1999. *Biochemistry (Mosc.) 51*, 3050–3058.

(48) Rockwell, N. C., Martin, S. S., Gulevich, A. G., and Lagarias, J. C. (2012) Phycoviolobilin Formation and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily. *Biochemistry (Mosc.) 51*, 1449–1463.

(49) Burgie, E. S., Walker, J. M., Phillips, G. N., and Vierstra, R. D. (2013) A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes. *Structure 21*, 88–97.

(50) Gottlieb, S. M., Kim, P. W., Corley, S. C., Madsen, D., Hanke, S. J., Chang, C.-W., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Larsen, D. S. (2014) Primary and Secondary Photodynamics of the Violet/Orange Dual-Cysteine NpF2164g3 Cyanobacteriochrome Domain from *Nostoc punctiforme*. *Biochemistry (Mosc.) 53*, 1029–1040.

(51) Narikawa, R., Enomoto, G., Ni-Ni-Win, Fushimi, K., and Ikeuchi, M. (2014) A New Type of Dual-Cys Cyanobacteriochrome GAF Domain Found in Cyanobacterium *Acaryochloris marina* , Which Has an Unusual Red/Blue Reversible Photoconversion Cycle. *Biochemistry (Mosc.) 53*, 5051–5059.

(52) Chen, Y., Zhang, J., Luo, J., Tu, J.-M., Zeng, X.-L., Xie, J., Zhou, M., Zhao, J.-Q., Scheer, H., and Zhao, K.-H. (2012) Photophysical diversity of two novel cyanobacteriochromes with phycocyanobilin chromophores: photochemistry and dark reversion kinetics. *FEBS J. 279*, 40–54.

(53) Christie, J. M., Arvai, A. S., Baxter, K. J., Heilmann, M., Pratt, A. J., O'Hara, A., Kelly, S. M., Hothorn, M., Smith, B. O., Hitomi, K., Jenkins, G. I., and Getzoff, E. D. (2012) Plant UVR8 Photoreceptor Senses UV-B by Tryptophan-Mediated Disruption of Cross-Dimer Salt Bridges. *Science 335*, 1492–1496.

(54) Di Wu, Hu, Q., Yan, Z., Chen, W., Yan, C., Huang, X., Zhang, J., Yang, P., Deng, H., Wang, J., Deng, X., and Shi, Y. (2012) Structural basis of ultraviolet-B perception by UVR8. *Nature 484*, 214–219.

(55) Ni, M., Tepperman, J. M., and Quail, P. H. (1999) Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature 400*, 781–784.

(56) Khanna, R., Huq, E., Kikis, E. A., Al-Sady, B., Lanzatella, C., and Quail, P. H. (2004) A Novel Molecular Recognition Motif Necessary for Targeting Photoactivated Phytochrome Signaling to Specific Basic Helix-Loop-Helix Transcription Factors. *Plant Cell 16*, 3033–3044.

(57) Liu, H., Yu, X., Li, K., Klejnot, J., Yang, H., Lisiero, D., and Lin, C. (2008) Photoexcited CRY2 Interacts with CIB1 to Regulate Transcription and Floral Initiation in Arabidopsis. *Science 322*, 1535–1539.

(58) Lee, S., Park, H., Kyung, T., Kim, N. Y., Kim, S., Kim, J., and Heo, W. D. (2014) Reversible protein inactivation by optogenetic trapping in cells. *Nat. Methods 11*, 633–636.

(59) Hellingwerf, K. J., Hendriks, J., and Gensch, T. (2003) Photoactive Yellow Protein, A New Type of Photoreceptor Protein: Will This "Yellow Lab" Bring Us Where We Want to Go?‖. *J. Phys. Chem. A 107*, 1082–1094.

(60) Ramachandran, P. L., Lovett, J. E., Carl, P. J., Cammarata, M., Lee, J. H., Jung, Y. O., Ihee, H., Timmel, C. R., and van Thor, J. J. (2011) The Short-Lived Signaling State of the Photoactive Yellow Protein Photoreceptor Revealed by Combined Structural Probes. *J. Am. Chem. Soc. 133*, 9395–9404.

(61) Nagel, G., Szellas, T., Huhn, W., Kateriya, S., Adeishvili, N., Berthold, P., Ollig, D., Hegemann, P., and Bamberg, E. (2003) Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl. Acad. Sci. 100*, 13940–13945.

(62) Ernst, O. P., Lodowski, D. T., Elstner, M., Hegemann, P., Brown, L. S., and Kandori, H. (2014) Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chem. Rev. 114*, 126–163.

(63) Möglich, A., and Moffat, K. (2007) Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. *J. Mol. Biol. 373*, 112–126.

(64) Müller, M., Bamann, C., Bamberg, E., and Kühlbrandt, W. (2011) Projection Structure of Channelrhodopsin-2 at 6 Å Resolution by Electron Crystallography. *J. Mol. Biol. 414*, 86–95.

(65) Kato, H. E., Zhang, F., Yizhar, O., Ramakrishnan, C., Nishizawa, T., Hirata, K., Ito, J., Aita, Y., Tsukazaki, T., Hayashi, S., Hegemann, P., Maturana, A. D., Ishitani, R., Deisseroth, K., and Nureki, O. (2012) Crystal structure of the channelrhodopsin light-gated cation channel. *Nature 482*, 369–374.

(66) Circolone, F., Granzin, J., Jentzsch, K., Drepper, T., Jaeger, K.-E., Willbold, D., Krauss, U., and Batra-Safferling, R. (2012) Structural Basis for the Slow Dark Recovery of a Full-Length LOV Protein from Pseudomonas putida. *J. Mol. Biol. 417*, 362–374.

(67) Diensthuber, R. P., Bommer, M., Gleichmann, T., and Möglich, A. (2013) Full-Length Structure of a Sensor Histidine Kinase Pinpoints Coaxial Coiled Coils as Signal Transducers and Modulators. *Structure 21*, 1127–1136.

(68) Takala, H., Björling, A., Berntsson, O., Lehtivuori, H., Niebling, S., Hoernke, M., Kosheleva, I., Henning, R., Menzel, A., Ihalainen, J. A., and Westenhoff, S. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature 509*, 245–248.

(69) Essen, L.-O., Mailliet, J., and Hughes, J. (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci. 105*, 14709–14714.

(70) Barends, T. R. M., Hartmann, E., Griese, J. J., Beitlich, T., Kirienko, N. V., Ryjenkov, D. A., Reinstein, J., Shoeman, R. L., Gomelsky, M., and Schlichting, I. (2009) Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature 459*, 1015–1018.

(71) Anantharaman, V., Balaji, S., and Aravind, L. (2006) The signaling helix: a common functional theme in diverse signaling proteins. *Biol. Direct 1*, 25.

(72) Möglich, A., Ayers, R. A., and Moffat, K. (2010) Addition at the molecular level: signal integration in designed Per-ARNT-Sim receptor proteins. *J. Mol. Biol. 400*, 477–486.

(73) Rockwell, N. C., Ohlendorf, R., and Möglich, A. (2013) Cyanobacteriochromes in full color and three dimensions. *Proc. Natl. Acad. Sci. 110*, 806–807.

(74) Huang, N., Chelliah, Y., Shan, Y., Taylor, C. A., Yoo, S.-H., Partch, C., Green, C. B., Zhang, H., and Takahashi, J. S. (2012) Crystal Structure of the Heterodimeric CLOCK:BMAL1 Transcriptional Activator Complex. *Science 337*, 189–194.

(75) Hulko, M., Berndt, F., Gruber, M., Linder, J. U., Truffault, V., Schultz, A., Martin, J., Schultz, J. E., Lupas, A. N., and Coles, M. (2006) The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell 126*, 929–940.

(76) Matthews, E. E., Zoonens, M., and Engelman, D. M. (2006) Dynamic helix interactions in transmembrane signaling. *Cell 127*, 447–450.

(77) Schultz, J. E., and Natarajan, J. (2013) Regulated unfolding: a basic principle of intraprotein signaling in modular proteins. *Trends Biochem. Sci. 38*, 538–545.

(78) Möglich, A., Ayers, R. A., and Moffat, K. (2009) Design and signaling mechanism of light-regulated histidine kinases. *J. Mol. Biol. 385*, 1433–1444.

(79) Pandit, J., Forman, M. D., Fennell, K. F., Dillman, K. S., and Menniti, F. S. (2009) Mechanism for the allosteric regulation of phosphodiesterase 2A deduced from the X-ray structure of a near full-length construct. *Proc. Natl. Acad. Sci. 106*, 18225–18230.

(80) Wang, C., Sang, J., Wang, J., Su, M., Downey, J. S., Wu, Q., Wang, S., Cai, Y., Xu, X., Wu, J., Senadheera, D. B., Cvitkovitch, D. G., Chen, L., Goodman, S. D., and Han, A. (2013) Mechanistic Insights Revealed by the Crystal Structure of a Histidine Kinase with Signal Transducer and Sensor Domains. *PLoS Biol.* (Stock, A., Ed.) *11*, e1001493.

(81) Aravind L, L. M. I. (2010) Natural History of Sensor Domains in bacterial signaling systems. *Sens. Mech. Bact. Mol. Asp. Signal Recognit.*

(82) Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J., and Punta, M. (2014) Pfam: the protein families database. *Nucleic Acids Res. 42*, D222–D230.

(83) Gao, R., and Stock, A. M. (2009) Biological Insights from Structures of Two-Component Proteins. *Annu. Rev. Microbiol. 63*, 133–154.

(84) Capra, E. J., and Laub, M. T. (2012) Evolution of Two-Component Signal Transduction Systems. *Annu. Rev. Microbiol. 66*, 325–347.

(85) Letunic, I., Doerks, T., and Bork, P. (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res. 43*, D257–D260.

(86) Stierl, M. (2013, September 30) Kleine Enzyme mit großer Perspektive. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I.

(87) Deisseroth, K., Feng, G., Majewska, A. K., Miesenböck, G., Ting, A., and Schnitzer, M. J. (2006) Next-Generation Optical Technologies for Illuminating Genetically Targeted Brain Circuits. *J. Neurosci. 26*, 10380–10386.

(88) Zemelman, B. V., Lee, G. A., Ng, M., and Miesenböck, G. (2002) Selective Photostimulation of Genetically ChARGed Neurons. *Neuron 33*, 15–22.

(89) Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., and Deisseroth, K. (2005) Millisecond-timescale, genetically targeted optical control of neural activity. *Nat. Neurosci. 8*, 1263–1268.

(90) Stierl, M., Stumpf, P., Udwari, D., Gueta, R., Hagedorn, R., Losi, A., Gartner, W., Petereit, L., Efetova, M., Schwarzel, M., Oertner, T. G., Nagel, G., and Hegemann, P. (2010) Light Modulation

of Cellular cAMP by a Small Bacterial Photoactivated Adenylyl Cyclase, bPAC, of the Soil Bacterium Beggiatoa. *J. Biol. Chem. 286*, 1181–1188.

(91) Jansen, V., Alvarez, L., Balbach, M., Strünker, T., Hegemann, P., Kaupp, U. B., and Wachten, D. (2015) Controlling fertilization and cAMP signaling in sperm by optogenetics. *eLife 4*, e05161.

(92) Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauß, U., Guterl, J.-K., Wendorff, M., Losi, A., Gärtner, W., and Jaeger, K.-E. (2007) Reporter proteins for in vivo fluorescence without oxygen. *Nat. Biotechnol. 25*, 443–445.

(93) Chapman, S., Faulkner, C., Kaiserli, E., Garcia-Mata, C., Savenkov, E. I., Roberts, A. G., Oparka, K. J., and Christie, J. M. (2008) The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection. *Proc. Natl. Acad. Sci. 105*, 20038–20043.

(94) Piatkevich, K. D., Subach, F. V., and Verkhusha, V. V. (2013) Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev. 42*, 3441.

(95) Shcherbakova, D. M., Shemetov, A. A., Kaberniuk, A. A., and Verkhusha, V. V. (2015) Natural Photoreceptors as a Source of Fluorescent Proteins, Biosensors, and Optogenetic Tools. *Annu. Rev. Biochem. 84*, 519–550.

(96) Kralj, J. M., Hochbaum, D. R., Douglass, A. D., and Cohen, A. E. (2011) Electrical Spiking in Escherichia coli Probed with a Fluorescent Voltage-Indicating Protein. *Science 333*, 345–348.

(97) Kralj, J. M., Douglass, A. D., Hochbaum, D. R., Maclaurin, D., and Cohen, A. E. (2011) Optical recording of action potentials in mammalian neurons using a microbial rhodopsin. *Nat. Methods 9*, 90–95.

(98) Gunaydin, L. A., Yizhar, O., Berndt, A., Sohal, V. S., Deisseroth, K., and Hegemann, P. (2010) Ultrafast optogenetic control. *Nat. Neurosci. 13*, 387–392.

(99) Berndt, A., Yizhar, O., Gunaydin, L. A., Hegemann, P., and Deisseroth, K. (2009) Bi-stable neural state switches. *Nat. Neurosci. 12*, 229–234.

(100) Bamann, C., Gueta, R., Kleinlogel, S., Nagel, G., and Bamberg, E. (2010) Structural Guidance of the Photocycle of Channelrhodopsin-2 by an Interhelical Hydrogen Bond. *Biochemistry (Mosc.) 49*, 267–278.

(101) Yizhar, O., Fenno, L. E., Prigge, M., Schneider, F., Davidson, T. J., O'Shea, D. J., Sohal, V. S., Goshen, I., Finkelstein, J., Paz, J. T., Stehfest, K., Fudim, R., Ramakrishnan, C., Huguenard, J. R., Hegemann, P., and Deisseroth, K. (2011) Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature 477*, 171–178.

(102) Kleinlogel, S., Feldbauer, K., Dempski, R. E., Fotis, H., Wood, P. G., Bamann, C., and Bamberg, E. (2011) Ultra light-sensitive and fast neuronal activation with the Ca2+-permeable channelrhodopsin CatCh. *Nat. Neurosci. 14*, 513–518.

(103) Wietek, J., Wiegert, J. S., Adeishvili, N., Schneider, F., Watanabe, H., Tsunoda, S. P., Vogt, A., Elstner, M., Oertner, T. G., and Hegemann, P. (2014) Conversion of Channelrhodopsin into a Light-Gated Chloride Channel. *Science 344*, 409–412.

(104) Berndt, A., Lee, S. Y., Ramakrishnan, C., and Deisseroth, K. (2014) Structure-Guided Transformation of Channelrhodopsin into a Light-Activated Chloride Channel. *Science 344*, 420–424.

(105) Lin, J. Y., Knutsen, P. M., Muller, A., Kleinfeld, D., and Tsien, R. Y. (2013) ReaChR: a red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation. *Nat. Neurosci. 16*, 1499–1508.

(106) Kim, J.-M., Hwa, J., Garriga, P., Reeves, P. J., RajBhandary, U. L., and Khorana, H. G. (2005) Light-Driven Activation of β2-Adrenergic Receptor Signaling by a Chimeric Rhodopsin Containing the β2-Adrenergic Receptor Cytoplasmic Loops†. *Biochemistry (Mosc.) 44*, 2284–2292.

(107) Airan, R. D., Thompson, K. R., Fenno, L. E., Bernstein, H., and Deisseroth, K. (2009) Temporally precise in vivo control of intracellular signalling. *Nature 458*, 1025–1029.

(108) Strickland, D., Moffat, K., and Sosnick, T. R. (2008) Light-activated DNA binding in a designed allosteric protein. *Proc. Natl. Acad. Sci. 105*, 10709–10714.

(109) Wu, Y. I., Frey, D., Lungu, O. I., Jaehrig, A., Schlichting, I., Kuhlman, B., and Hahn, K. M. (2009) A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature 461*, 104–108.

(110) Pham, E., Mills, E., and Truong, K. (2011) A Synthetic Photoactivated Protein to Generate Local or Global Ca2+ Signals. *Chem. Biol. 18*, 880–890.

(111) Mills, E., Chen, X., Pham, E., Wong, S., and Truong, K. (2012) Engineering a Photoactivated Caspase-7 for Rapid Induction of Apoptosis. *ACS Synth. Biol. 1*, 75–82.

(112) Schmidt, D., Tillberg, P. W., Chen, F., and Boyden, E. S. (2014) A fully genetically encoded protein architecture for optical control of peptide ligand concentration. *Nat. Commun. 5*.

(113) Renicke, C., Schuster, D., Usherenko, S., Essen, L.-O., and Taxis, C. (2013) A LOV2 Domain-Based Optogenetic Tool to Control Protein Degradation and Cellular Function. *Chem. Biol. 20*, 619–626.

(114) Bonger, K. M., Rakhit, R., Payumo, A. Y., Chen, J. K., and Wandless, T. J. (2014) General Method for Regulating Protein Stability with Light. *ACS Chem. Biol. 9*, 111–115.

(115) Yi, J. J., Wang, H., Vilela, M., Danuser, G., and Hahn, K. M. (2014) Manipulation of Endogenous Kinase Activity in Living Cells Using Photoswitchable Inhibitory Peptides. *ACS Synth. Biol. 3*, 788–795.

(116) Niopek, D., Benzinger, D., Roensch, J., Draebing, T., Wehler, P., Eils, R., and Di Ventura, B. (2014) Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells. *Nat. Commun. 5*, 4404.

(117) Strickland, D., Lin, Y., Wagner, E., Hope, C. M., Zayner, J., Antoniou, C., Sosnick, T. R., Weiss, E. L., and Glotzer, M. (2012) TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods 9*, 379–384.

(118) van Bergeijk, P., Adrian, M., Hoogenraad, C. C., and Kapitein, L. C. (2015) Optogenetic control of organelle transport and positioning. *Nature 518*, 111–114.

(119) Lungu, O. I., Hallett, R. A., Choi, E. J., Aiken, M. J., Hahn, K. M., and Kuhlman, B. (2012) Designing Photoswitchable Peptides Using the AsLOV2 Domain. *Chem. Biol. 19*, 507–517.

(120) Guntas, G., Hallett, R. A., Zimmerman, S. P., Williams, T., Yumerefendi, H., Bear, J. E., and Kuhlman, B. (2015) Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc. Natl. Acad. Sci. 112*, 112–117.

(121) Lee, J., Natarajan, M., Nashine, V. C., Socolich, M., Vo, T., Russ, W. P., Benkovic, S. J., and Ranganathan, R. (2008) Surface Sites for Engineering Allosteric Control in Proteins. *Science 322*, 438–442.

(122) Wang, H. H., Kim, H., Cong, L., Jeong, J., Bang, D., and Church, G. M. (2012) Genome-scale promoter engineering by coselection MAGE. *Nat. Methods 9*, 591–593.

(123) Nihongaki, Y., Suzuki, H., Kawano, F., and Sato, M. (2014) Genetically Engineered Photoinducible Homodimerization System with Improved Dimer-Forming Efficiency. *ACS Chem. Biol. 9*, 617–621.

(124) Kawano, F., Suzuki, H., Furuya, A., and Sato, M. (2015) Engineered pairs of distinct photoswitches for optogenetic control of cellular proteins. *Nat. Commun. 6*, 6256.

(125) Shimizu-Sato, S., Huq, E., Tepperman, J. M., and Quail, P. H. (2002) A light-switchable gene promoter system. *Nat. Biotechnol. 20*, 1041–1044.

(126) Strauss, H. M., Schmieder, P., and Hughes, J. (2005) Light-dependent dimerisation in the N-terminal sensory module of cyanobacterial phytochrome 1. *FEBS Lett. 579*, 3970–3974.

(127) Tyszkiewicz, A. B., and Muir, T. W. (2008) Activation of protein splicing with light in yeast. *Nat. Methods*.

(128) Levskaya, A., Weiner, O. D., Lim, W. A., and Voigt, C. A. (2009) Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature 461*, 997–1001.

(129) Müller, K., Engesser, R., Schulz, S., Steinberg, T., Tomakidi, P., Weber, C. C., Ulm, R., Timmer, J., Zurbriggen, M. D., and Weber, W. (2013) Multi-chromatic control of mammalian gene expression and signaling. *Nucleic Acids Res. 41*, e124–e124.

(130) Toettcher, J. E., Weiner, O. D., and Lim, W. A. (2013) Using Optogenetics to Interrogate the Dynamic Control of Signal Transmission by the Ras/Erk Module. *Cell 155*, 1422–1434.

(131) Yang, X., Jost, A. P.-T., Weiner, O. D., and Tang, C. (2013) A light-inducible organelle-targeting system for dynamically activating and inactivating signaling in budding yeast. *Mol. Biol. Cell 24*, 2419–2430.

(132) Beyer, H. M., Juillot, S., Herbst, K., Samodelov, S. L., Müller, K., Schamel, W. W., Römer, W., Schäfer, E., Nagy, F., Strähle, U., Weber, W., and Zurbriggen, M. D. (2015) Red Light-Regulated Reversible Nuclear Localization of Proteins in Mammalian Cells and Zebrafish. *ACS Synth. Biol. 4*, 951–958.

(133) Bugaj, L. J., Choksi, A. T., Mesuda, C. K., Kane, R. S., and Schaffer, D. V. (2013) Optogenetic protein clustering and signaling activation in mammalian cells. *Nat. Methods 10*, 249–252.

(134) Chang, K.-Y., Woo, D., Jung, H., Lee, S., Kim, S., Won, J., Kyung, T., Park, H., Kim, N., Yang, H. W., Park, J.-Y., Hwang, E. M., Kim, D., and Do Heo, W. (2014) Light-inducible receptor tyrosine kinases that regulate neurotrophin signalling. *Nat. Commun. 5*.

(135) Kim, N., Kim, J. M., Lee, M., Kim, C. Y., Chang, K.-Y., and Heo, W. D. (2014) Spatiotemporal Control of Fibroblast Growth Factor Receptor Signals by Blue Light. *Chem. Biol. 21*, 903–912.

(136) Wend, S., Wagner, H. J., Müller, K., Zurbriggen, M. D., Weber, W., and Radziwill, G. (2014) Optogenetic Control of Protein Kinase Activity in Mammalian Cells. *ACS Synth. Biol. 3*, 280–285.

(137) Taslimi, A., Vrana, J. D., Chen, D., Borinskaya, S., Mayer, B. J., Kennedy, M. J., and Tucker, C. L. (2014) An optimized optogenetic clustering tool for probing protein interaction and function. *Nat. Commun. 5*, 4925.

(138) Kennedy, M. J., Hughes, R. M., Peteya, L. A., Schwartz, J. W., Ehlers, M. D., and Tucker, C. L. (2010) Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods 7*, 973–975.

(139) Liu, H., Gomez, G., Lin, S., Lin, S., and Lin, C. (2012) Optogenetic Control of Transcription in Zebrafish. *PLoS ONE* (Chatterji, D., Ed.) *7*, e50738.

(140) Idevall-Hagren, O., Dickson, E. J., Hille, B., Toomre, D. K., and Camilli, P. D. (2012) Optogenetic control of phosphoinositide metabolism. *Proc. Natl. Acad. Sci. 109*, 13894–13895.

(141) Kakumoto, T., and Nakata, T. (2013) Optogenetic Control of PIP3: PIP3 Is Sufficient to Induce the Actin-Based Active Part of Growth Cones and Is Regulated via Endocytosis. *PLoS ONE* (Datta, P. K., Ed.) *8*, e70861.

(142) Aoki, K., Kumagai, Y., Sakurai, A., Komatsu, N., Fujita, Y., Shionyu, C., and Matsuda, M. (2013) Stochastic ERK Activation Induced by Noise and Cell-to-Cell Propagation Regulates Cell Density-Dependent Proliferation. *Mol. Cell 52*, 529–540.

(143) Zhang, K., Duan, L., Ong, Q., Lin, Z., Varman, P. M., Sung, K., and Cui, B. (2014) Light-Mediated Kinetic Control Reveals the Temporal Effect of the Raf/MEK/ERK Pathway in PC12 Cell Neurite Outgrowth. *PLoS ONE 9*, e92917.

(144) Konermann, S., Brigham, M. D., Trevino, A. E., Hsu, P. D., Heidenreich, M., Le Cong, Platt, R. J., Scott, D. A., Church, G. M., and Zhang, F. (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature 500*, 472–476.

(145) Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H., and Sato, M. (2015) CRISPR-Cas9-based Photoactivatable Transcription System. *Chem. Biol. 22*, 169–174.

(146) Polstein, L. R., and Gersbach, C. A. (2015) A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol. 11*, 198–200.

(147) Chen, D., Gibson, E. S., and Kennedy, M. J. (2013) A light-triggered protein secretion system. *J. Cell Biol. 201*, 631–640.

(148) Crefcoeur, R. P., Yin, R., Ulm, R., and Halazonetis, T. D. (2013) Ultraviolet-B-mediated induction of protein–protein interactions in mammalian cells. *Nat. Commun. 4*, 1779.

(149) Zhou, X. X., Chung, H. K., Lam, A. J., and Lin, M. Z. (2012) Optical Control of Protein Activity by Fluorescent Protein Domains. *Science 338*, 810–814.

(150) Yazawa, M., Sadaghiani, A. M., Hsueh, B., and Dolmetsch, R. E. (2009) Induction of protein-protein interactions in live cells using light. *Nat. Biotechnol. 27*, 941–945.

(151) Polstein, L. R., and Gersbach, C. A. (2012) Light-Inducible Spatiotemporal Control of Gene Activation by Customizable Zinc Finger Transcription Factors. *J. Am. Chem. Soc. 134*, 16480–16483.

(152) Motta-Mena, L. B., Reade, A., Mallory, M. J., Glantz, S., Weiner, O. D., Lynch, K. W., and Gardner, K. H. (2014) An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nat. Chem. Biol.*

(153) Grusch, M., Schelch, K., Riedler, R., Reichhart, E., Differ, C., Berger, W., Ingles-Prieto, A., and Janovjak, H. (2014) Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. *EMBO J. 33*, 1713–1726.

(154) Dixon, R. E., Yuan, C., Cheng, E. P., Navedo, M. F., and Santana, L. F. (2012) Ca2+ signaling amplification by oligomerization of L-type Cav1.2 channels. *Proc. Natl. Acad. Sci. 109*, 1749–1754.

(155) Hawking, S. W. (2001) The Universe in a Nutshell 1st ed. Bantam.

(156) Levskaya, A., Chevalier, A. A., Tabor, J. J., Simpson, Z. B., Lavery, L. A., Levy, M., Davidson, E. A., Scouras, A., Ellington, A. D., Marcotte, E. M., and Voigt, C. A. (2005) Synthetic biology: engineering Escherichia coli to see light. *Nature 438*, 441–442.

(157) Ryu, M.-H., Kang, I.-H., Nelson, M. D., Jensen, T. M., Lyuksyutova, A. I., Siltberg-Liberles, J., Raizen, D. M., and Gomelsky, M. (2014) Engineering adenylate cyclases regulated by near-infrared window light. *Proc. Natl. Acad. Sci. 111*, 10167–10172.

(158) Gasser, C., Taiber, S., Yeh, C.-M., Wittig, C. H., Hegemann, P., Ryu, S., Wunder, F., and Moglich, A. (2014) Engineering of a red-light-activated human cAMP/cGMP-specific phosphodiesterase. *Proc. Natl. Acad. Sci. 111*, 8803–8808.

(159) Ohlendorf, R., Vidavski, R. R., Eldar, A., Moffat, K., and Möglich, A. (2012) From Dusk till Dawn: One-Plasmid Systems for Light-Regulated Gene Expression. *J. Mol. Biol. 416*, 534–542.

(160) Gleichmann, T., Diensthuber, R. P., and Möglich, A. (2013) Charting the Signal Trajectory in a Light-Oxygen-Voltage Photoreceptor by Random Mutagenesis and Covariance Analysis. *J. Biol. Chem. 288*, 29345–29355.

(161) Diensthuber, R. P., Engelhard, C., Lemke, N., Gleichmann, T., Ohlendorf, R., Bittl, R., and Möglich, A. (2014) Biophysical, Mutational, and Functional Investigation of the Chromophore-Binding Pocket of Light-Oxygen-Voltage Photoreceptors. *ACS Synth. Biol. 3*, 811–819.

(162) Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343–345.

(163) Strack, R. L., Strongin, D. E., Bhattacharyya, D., Tao, W., Berman, A., Broxmeyer, H. E., Keenan, R. J., and Glick, B. S. (2008) A noncytotoxic DsRed variant for whole-cell labeling. *Nat. Methods 5*, 955–957.

(164) Ostermeier, M., and Lutz, S. (2003) The creation of ITCHY hybrid protein libraries. *Methods Mol. Biol. Clifton NJ 231*, 129–141.

(165) Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/green cyanobacteriochromes: Sensors of color and power. *Biochemistry (Mosc.) 51*, 9667–9677.

(166) Consortium, T. U. (2015) UniProt: a hub for protein information. *Nucleic Acids Res. 43*, D204–D212.

(167) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinforma. Oxf. Engl. 23*, 2947–2948.

(168) Gambetta, G. A., and Lagarias, J. C. (2001) Genetic engineering of phytochrome biosynthesis in bacteria. *Proc. Natl. Acad. Sci. 98*, 10566–10571.

(169) Mukougawa, K., Kanamoto, H., Kobayashi, T., Yokota, A., and Kohchi, T. (2006) Metabolic engineering to produce phytochromes with phytochromobilin, phycocyanobilin, or phycoerythrobilin chromophore in Escherichia coli. *FEBS Lett. 580*, 1333–1338.

(170) Calos, M. P. (1978) DNA sequence for a low-level promoter of the lac repressor gene and an "up" promoter mutation. *Nature 274*, 762–765.

(171) Hirose, Y., Shimada, T., Narikawa, R., Katayama, M., and Ikeuchi, M. (2008) Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phycobilisome linker protein. *Proc. Natl. Acad. Sci. 105*, 9528–9533.

(172) Kasahara, M., Swartz, T. E., Olney, M. A., Onodera, A., Mochizuki, N., Fukuzawa, H., Asamizu, E., Tabata, S., Kanegae, H., Takano, M., Christie, J. M., Nagatani, A., and Briggs, W. R. (2002) Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from Arabidopsis, rice, and Chlamydomonas reinhardtii. *Plant Physiol. 129*, 762–773.

(173) Linder, J. U., and Schultz, J. E. (2003) The class III adenylyl cyclases: multi-purpose signalling modules. *Cell. Signal. 15*, 1081–1089.

(174) Miller, J. H. (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press,U.S.

(175) Dickson, R. C., Abelson, J., Barnes, W. M., and Reznikoff, W. S. (1975) Genetic regulation: the Lac control region. *Science 187*, 27–35.

(176) Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci. 95*, 5752–5756.

(177) Kanacher, T., Schultz, A., Linder, J. U., and Schultz, J. E. (2002) A GAF-domain-regulated adenylyl cyclase from Anabaena is a self-activating cAMP switch. *EMBO J. 21*, 3672–3680.

(178) Bruder, S., Linder, J. U., Martinez, S. E., Zheng, N., Beavo, J. A., and Schultz, J. E. (2005) The cyanobacterial tandem GAF domains from the cyaB2 adenylyl cyclase signal via both cAMP-binding sites. *Proc. Natl. Acad. Sci. U. S. A. 102*, 3088–3092.

(179) Schmidl, S. R., Sheth, R. U., Wu, A., and Tabor, J. J. (2014) Refactoring and Optimization of Light-Switchable *Escherichia coli* Two-Component Systems. *ACS Synth. Biol. 3*, 820–831.

(180) Winkler, K., Schultz, A., and Schultz, J. E. (2012) The S-Helix Determines the Signal in a Tsr Receptor/Adenylyl Cyclase Reporter. *J. Biol. Chem. 287*, 15479–15488.

(181) Ostermeier, M., Shim, J. H., and Benkovic, S. J. (1999) A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat. Biotechnol. 17*, 1205–1209.

(182) Möglich, A., Ayers, R. A., and Moffat, K. (2009) Structure and signaling mechanism of Per-ARNT-Sim domains. *Struct. Lond. Engl. 1993 17*, 1282–1294.

(183) Ikeuchi, M., and Ishizuka, T. (2008) Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. *Photochem. Photobiol. Sci. 7*, 1159–1167.

(184) Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/Green Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry (Mosc.) 51*, 9667–9677.

(185) Tabor, J. J., Levskaya, A., and Voigt, C. A. (2011) Multichromatic control of gene expression in Escherichia coli. *J. Mol. Biol. 405*, 315–324.

(186) Olson, E. J., Hartsough, L. A., Landry, B. P., Shroff, R., and Tabor, J. J. (2014) Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nat. Methods 11*, 449–455.

(187) Ostermeier, M., Nixon, A. E., Shim, J. H., and Benkovic, S. J. (1999) Combinatorial protein engineering by incremental truncation. *Proc. Natl. Acad. Sci. 96*, 3562–3567.

(188) Stefan Lutz, Marc Ostermeier, and Stephen J Benkovic. Rapid generation of incremental truncation libraries for protein engineering using α-phosphothioate nucleotides.

(189) Sieber, V., Martinez, C. A., and Arnold, F. H. (2001) Libraries of hybrid proteins from distantly related sequences. *Nat. Biotechnol. 19*, 456–460.

(190) Ostermeier, M. (2003) Theoretical distribution of truncation lengths in incremental truncation libraries. *Biotechnol. Bioeng. 82*, 564–577.

(191) Quan, J., and Tian, J. (2009) Circular polymerase extension cloning of complex gene libraries and pathways. *PloS One 4*, e6441.

(192) Möglich, A., Ayers, R. A., and Moffat, K. (2009) Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Structure 17*, 1282–1294.

(193) Herrou, J., and Crosson, S. (2011) Function, structure and mechanism of bacterial photosensory LOV proteins. *Nat. Rev. Microbiol. 9*, 713–723.

(194) Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*.

(195) Wolgemuth, C. W., and Sun, S. X. (2006) Elasticity of α -Helical Coiled Coils. *Phys. Rev. Lett. 97*.

(196) Mason, J. M., and Arndt, K. M. (2004) Coiled Coil Domains: Stability, Specificity, and Biological Implications. *ChemBioChem 5*, 170–176.

(197) Lupas, A. N., and Gruber, M. (2005) The Structure of α-Helical Coiled Coils (Chemistry, B.-A. in P., Ed.), pp 37–38. Academic Press.

(198) Lupas, A., Dyke, M. V., and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science 252*, 1162–1164.

(199) Raffelberg, S., Wang, L., Gao, S., Losi, A., Gärtner, W., and Nagel, G. (2013) A LOV-domain-mediated blue-light-activated adenylate (adenylyl) cyclase from the cyanobacterium *Microcoleus chthonoplastes* PCC 7420. *Biochem. J. 455*, 359–365.

(200) Hulko, M., Berndt, F., Gruber, M., Linder, J. U., Truffault, V., Schultz, A., Martin, J., Schultz, J. E., Lupas, A. N., and Coles, M. (2006) The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell 126*, 929–940.

(201) Ferris, H. U., Dunin-Horkawicz, S., Mondéjar, L. G., Hulko, M., Hantke, K., Martin, J., Schultz, J. E., Zeth, K., Lupas, A. N., and Coles, M. (2011) The Mechanisms of HAMP-Mediated Signaling in Transmembrane Receptors. *Structure 19*, 378–385.

(202) Ferris, H. U., Dunin-Horkawicz, S., Hornig, N., Hulko, M., Martin, J., Schultz, J. E., Zeth, K., Lupas, A. N., and Coles, M. (2012) Mechanism of Regulation of Receptor Histidine Kinases. *Structure 20*, 56–66.

(203) Ferris, H. U., Zeth, K., Hulko, M., Dunin-Horkawicz, S., and Lupas, A. N. (2014) Axial helix rotation as a mechanism for signal regulation inferred from the crystallographic analysis of the E. coli serine chemoreceptor. *J. Struct. Biol. 186*, 349–356.

(204) Wang, B., Zhao, A., Novick, R. P., and Muir, T. W. (2014) Activation and Inhibition of the Receptor Histidine Kinase AgrC Occurs through Opposite Helical Transduction Motions. *Mol. Cell 53*, 929–940.

(205) Dago, A. E., Schug, A., Procaccini, A., Hoch, J. A., Weigt, M., and Szurmant, H. (2012) Structural basis of histidine kinase autophosphorylation deduced by integrating genomics, molecular dynamics, and mutagenesis. *Proc. Natl. Acad. Sci. 109*, E1733–E1742.

(206) Cheung, J., and Hendrickson, W. A. (2009) Structural Analysis of Ligand Stimulation of the Histidine Kinase NarX. *Structure 17*, 190–201.

(207) Takala, H., Björling, A., Linna, M., Westenhoff, S., and Ihalainen, J. A. (2015) Light-induced Changes in the Dimerization Interface of Bacteriophytochromes. *J. Biol. Chem. 290*, 16383–16392.

(208) Björling, A., Berntsson, O., Takala, H., Gallagher, K. D., Patel, H., Gustavsson, E., St. Peter, R., Duong, P., Nugent, A., Zhang, F., Berntsen, P., Appio, R., Rajkovic, I., Lehtivuori, H., Panman, M. R., Hoernke, M., Niebling, S., Harimoorthy, R., Lamparter, T., Stojković, E. A., Ihalainen, J. A., and Westenhoff, S. (2015) Ubiquitous Structural Signaling in Bacterial Phytochromes. *J. Phys. Chem. Lett. 6*, 3379–3383.

(209) Burgie, E. S., Wang, T., Bussell, A. N., Walker, J. M., Li, H., and Vierstra, R. D. (2014) Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements during Photoconversion. *J. Biol. Chem. 289*, 24573–24587.

(210) Monod, J., Wyman, J., and Changeux, J.-P. (1965) On the nature of allosteric transitions: A plausible model. *J. Mol. Biol. 12*, 88–118.

(211) Kanchan, K., Linder, J., Winkler, K., Hantke, K., Schultz, A., and Schultz, J. E. (2010) Transmembrane Signaling in Chimeras of the Escherichia coli Aspartate and Serine Chemotaxis Receptors and Bacterial Class III Adenylyl Cyclases. *J. Biol. Chem. 285*, 2090–2099.

(212) Tesmer, J. J., Sunahara, R. K., Gilman, A. G., and Sprang, S. R. (1997) Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsα· GTPγS. *Science 278*, 1907–1916.

(213) Tews, I., Findeisen, F., Sinning, I., Schultz, A., Schultz, J. E., and Linder, J. U. (2005) The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme. *Science 308*, 1020–1023.

(214) Topal, H., Fulcher, N. B., Bitterman, J., Salazar, E., Buck, J., Levin, L. R., Cann, M. J., Wolfgang, M. C., and Steegborn, C. (2012) Crystal Structure and Regulation Mechanisms of the CyaB Adenylyl Cyclase from the Human Pathogen Pseudomonas aeruginosa. *J. Mol. Biol. 416*, 271–286.

(215) Zähringer, F., Lacanna, E., Jenal, U., Schirmer, T., and Boehm, A. (2013) Structure and Signaling Mechanism of a Zinc-Sensory Diguanylate Cyclase. *Structure 21*, 1149–1157.

(216) Klapoetke, N. C., Murata, Y., Kim, S. S., Pulver, S. R., Birdsey-Benson, A., Cho, Y. K., Morimoto, T. K., Chuong, A. S., Carpenter, E. J., Tian, Z., Wang, J., Xie, Y., Yan, Z., Zhang, Y., Chow, B. Y., Surek, B., Melkonian, M., Jayaraman, V., Constantine-Paton, M., Wong, G. K.-S., and Boyden, E. S. (2014) Independent optical excitation of distinct neural populations. *Nat. Methods 11*, 338–346.

(217) Hell, S. W., and Wichmann, J. (1994) Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett. 19*, 780–782.

(218) Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian Expression of Infrared Fluorescent Proteins Engineered from a Bacterial Phytochrome. *Science 324*, 804–807.

(219) Filonov, G. S., Piatkevich, K. D., Ting, L.-M., Zhang, J., Kim, K., and Verkhusha, V. V. (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol. 29*, 757–761.

(220) Narikawa, R., Nakajima, T., Aono, Y., Fushimi, K., Enomoto, G., Ni-Ni-Win, Itoh, S., Sato, M., and Ikeuchi, M. (2015) A biliverdin-binding cyanobacteriochrome from the chlorophyll d–bearing cyanobacterium Acaryochloris marina. *Sci. Rep. 5*, 7950.

(221) Ryu, M.-H., and Gomelsky, M. (2014) Near-infrared Light Responsive Synthetic c-di-GMP Module for Optogenetic Applications. *ACS Synth. Biol. 3*, 802–810.

(222) Müller, K., Engesser, R., Timmer, J., Nagy, F., Zurbriggen, M. D., and Weber, W. (2013) Synthesis of phycocyanobilin in mammalian cells. *Chem. Commun. 49*, 8970.

(223) Scheib, U., Stehfest, K., Gee, C. E., Körschen, H. G., Fudim, R., Oertner, T. G., and Hegemann, P. (2015) The rhodopsin–guanylyl cyclase of the aquatic fungus Blastocladiella emersonii enables fast optical control of cGMP signaling. *Sci Signal 8*, rs8–rs8.

(224) Luck, M., Mathes, T., Bruun, S., Fudim, R., Hagedorn, R., Tran Nguyen, T. M., Kateriya, S., Kennis, J. T. M., Hildebrandt, P., and Hegemann, P. (2012) A photochromic histidine kinase

rhodopsin (HKR1) that is bimodally switched by ultraviolet and blue light. *J. Biol. Chem. 287*, 40083–40090.

(225) Schumacher, C. H., Körschen, H. G., Nicol, C., Gasser, C., Seifert, R., Schwärzel, M., and Möglich, A. A Fluorometric Activity Assay for Light-Regulated Cyclic-Nucleotide-Monophosphate Actuators (Kianianmomeni, A., Ed.). Springer New York.

(226) Xu, L., Jin, X., Rainey, G. J., Wu, H., and Gao, C. (2013) A mammalian expression system for high throughput antibody screening. *J. Immunol. Methods 395*, 45–53.

(227) Henry, J. T., and Crosson, S. (2011) Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context. *Annu. Rev. Microbiol. 65*, 261–286.

(228) Francis, S. H., Blount, M. A., and Corbin, J. D. (2011) Mammalian Cyclic Nucleotide Phosphodiesterases: Molecular Mechanisms and Physiological Functions. *Physiol. Rev. 91*, 651–690.

8 Appendix

8.1 Abbreviations

aa	amino acid
Å	Ångström
Amp	ampicillin
*As*LOV2	phototropin LOV2 domain from *Avena sativa*
*At*Cry2	cryptochrome 2 from *Arabidopsis thaliana*
ATP	adenosine trisphosphate
BPhy	bacteriophytochrome
BLUF	sensors of blue light using flavin adenine dinucleotide
C-terminus	carboxy terminus
CA	catalytic domain
cAMP	cyclic adenosine monophosphate
CAP	catabolite activator protein
	(also known as: cAMP receptor protein - CRP)
CBCR	cyanobacteriochrome
cGMP	cyclic guanosine mononucleotide
CIB1	cryptochrome 2-interacting basic helix-loop-helix 1
CYCc	adenylate/guanylate cyclase
DHp	dimerization and histidine phosphotransfer
DIT	aspartate-isoleucine-threonine
DNA	deoxyribonucleic acid
dNTP	deooxynucleotide
*Dr*BPhy	bacteriophytochrome from *Deinococcus radiodurans*
ε	extinction coefficient
FACS	fluorescence-activated cell sorting
FAD	flavin-adenine dinucleotide
FMN	flavin mononucleotide
FP	fluorescent protein
GAF	cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA

GFP	green-fluorescent protein
GGDEF	diguanylate cyclase
GPCR	G-protein coupled receptor
G/R	green-red photocycle
GTP	guanosine trisphosphate
HAMP	histidine kinase, adenylate cyclase, methyl accepting protein and phosphatases
HCl	hydrogen chloride
His	histidine
HisK	histidine kinase
HPLC	high-performance liquid chromatography
LB	lysogeny broth
LED	light-emitting diode
IPTG	isopropyl- β -d-thiogalactopyranoside
ITCHY	incremental truncation for the creation of hybrid enzymes
Kan	kanamycin
kb	kilo bases
kDa	kilo Dalton
LOV	light oxygen voltage
M	molar (mol l^{-1})
MCP	methyl-accepting chemotaxis protein
MCS	multiple-cloning site
$MgCl_2$	magnesium chloride
N-terminus	amino terminus
NaCl	sodium chloride
NAD	nicotinamide adenine dinucleotide
OD_{600}	optical density at 600 nm
P(r/fr/b/g)	photostate absorbing red/far-red/blue/green light
PAGE	poly-acrylamide gel electrophoresis
PAS	Per-ARNT-Sim
PATCHY	primer-aided truncation for the creation of hybrid enzymes
PCB	phycocyanobilin

PCR	polymerase chain reaction
PDE	phosphodiesterase
PEG	polyethylene glycol
PVB	phycoviolobilin
PYP	photoactive yellow protein
R/G	red-green photocycle
RR	response regulator
SDS	sodium dodecyl sulfate
SHK	sensor histidine kinase
STAS	sulfate transporter/anti-sigma-factor antagonist
Strep	streptomycin
TCS	two-component system
Tris	Tris(hydroxymethyl)aminomethane
UV(-B)	ultra violet (-B)
UVR8	UV-B resistance 8

8.2 List of publications

Ohlendorf R, Schumacher CH, Richter F, Möglich A. Library-aided probing of linker determinants in hybrid photoreceptors. ACS Synthetic Biology (accepted)

Diensthuber RP, Engelhard C, Lemke N, Gleichmann T, **Ohlendorf R**, Bittl R, Möglich A. (2014) Biophysical, mutational and functional investigation of the chromophore-binding pocket of light-oxygen-voltage photoreceptors. ACS Synthetic Biology 3(11), 811-819

Diensthuber RP, **Ohlendorf R**, Gleichmann T, Schubert R, Möglich A. (2013) Lichtregulierte Genexpression. BioSpektrum 19(2), 149-151

Rockwell NC, **Ohlendorf R**, Möglich A. (2013) Cyanobacteriochromes in full color and three dimensions. Proceedings of the National Academy of Sciences U S A 110(3), 806-7

Ohlendorf R, Vidavski RR, Eldar A, Moffat K, Möglich A. (2012) From Dusk Till Dawn: One-Plasmid Systems for Light-Regulated Gene Expression. Journal of Molecular Biology (416), 534-542

8.3 Symposia and meeting contributions

03/2015 **Ohlendorf R**, Möglich A. A library approach for generating light-switchable fusion proteins. Keystone Symposium on Optogenetics (Denver, CO, USA), Poster

05/2014 **Ohlendorf R**, Möglich A. High-throughput approach for generating light-switchable fusion proteins. Gordon Research Conference & Seminar "Photosensory Receptors & Signal Transduction", (Barga, Italy), Poster

07/2013 **Ohlendorf R**, Moffat K, Möglich A. From Dusk till Dawn: One-Plasmid Systems for Light-Regulated Gene Expression. The sixth international meeting on synthetic biology (SB6.0), Biobricks Foundation, Imperial College London (London, UK), Poster

11/2011 **Ohlendorf R**, Moffat K, Möglich A. From Dusk till Dawn: One-Plasmid Systems for Light-Regulated Gene Expression. Janelia Farm Conference (HHMI) „Single Molecules Meet Systems Biology" (Ashburn, VA, USA), Poster

8.4 Eigenständigkeitserklärung

Hiermit versichere ich, dass ich auf Grundlage der angegebenen Hilfen und Hilfsmittel die Dissertation selbstständig erarbeitet und verfasst habe.

Berlin, den

_____

Robert Ohlendorf

8.5 Acknowledgments

Ich möchte zuerst **Prof. Dr. Andreas Möglich** für die lehrreiche Zeit, ausgiebige wissenschaftliche Diskussionen und die jederzeit offene Tür danken.

Zudem geht ein großer Dank an **Carlos, Roman, Florian, Ralph** und die ganze **AG Möglich** und **AG Hegemann** für eine Arbeitsatmosphäre, wie man sie sich besser nicht wünschen kann. Danke an **Charlotte** für ihren Einsatz bei der Vorbereitung der NGS-Proben, sowie **Andreas** für die Datenanalyse.
Für wissenschaftlichen Rat und diverse Materialien möchte ich **Dr. Anke Licht** (HU Berlin) und **Dr. Nathan Rockwell** (UC Davis) danken.

Hervorheben möchte ich **Johannes** und **Thea**, die nicht nur gute Kollegen, sondern auch großartige Freunde sind. Gleiches gilt für **Franzi**, die zudem durch hilfreiche Kommentare zu dieser Arbeit beigetragen hat. Danke euch für die gute Zeit!

Besonderen Dank an **Chrissi**, nicht nur für das Lesen und deinen Rat zu dieser Arbeit, sondern für einen Rückhalt, der nicht hoch genug einzuschätzen ist.

An letzter, aber eigentlich erster Stelle möchte ich **meinen Eltern und der ganzen Familie** danken, die mich nun seit fast dreißig Jahren in meinem Tun bestärken.